

THE EXPURGATION-AUGMENTATION METHOD FOR CONSTRUCTING GOOD PLANE SUBSPACE CODES

JINGMEI AI, THOMAS HONOLD, AND HAITENG LIU

Department of Information Science and Electronics Engineering
Zhejiang University, 38 Zheda Road, 310027 Hangzhou, China

ABSTRACT. As shown in [31], one of the five isomorphism types of optimal binary subspace codes of size 77 for packet length $v = 6$, constant dimension $k = 3$ and minimum subspace distance $d = 4$ can be constructed by first expurgating and then augmenting the corresponding lifted Gabidulin code in a fairly simple way. The method was refined in [36, 29] to yield an essentially computer-free construction of a currently best-known plane subspace code of size 329 for $(v, k, d) = (7, 3, 4)$. In this paper we generalize the expurgation-augmentation approach to arbitrary packet length v , providing both a detailed theoretical analysis of our method and computational results for small parameters. As it turns out, our method is capable of producing codes larger than those obtained by the echelon-Ferrers construction and its variants. We are able to prove this observation rigorously for packet lengths $v \equiv 3 \pmod{4}$.

1. INTRODUCTION

Let V be a vector space of finite dimension v over the finite field \mathbb{F}_q . The lattice $\text{PG}(V)$ of all subspaces of V , relative to the operations $X \wedge Y = X \cap Y$ (meet) and $X \vee Y = X + Y$ (join) is called the projective (coordinate) geometry associated with V and forms the unique (up to isomorphism) model of $\text{PG}(v-1, \mathbb{F}_q)$, the Desarguesian projective geometry of geometric dimension $v-1$ and order q . Moreover, $\text{PG}(V)$ forms a metric space with respect to the *subspace distance* defined by $d_s(X, Y) = d_s(X+Y) - d_s(X \cap Y) = \dim(X) + \dim(Y) - 2\dim(X \cap Y)$. A code in this metric space is called a *q -ary subspace code*.¹ Such codes are of interest in the emerging area of error-resilient network coding, where they can be used as channel codes for the linear operator channel introduced by Koetter and Kschischang [32]; cf. also [34, 26, 23]. The so-called *Main Problem of Subspace Coding*, motivated by the application to network coding and modelled after the Main Problem of classical coding theory, asks for the determination (and, as a secondary goal, the classification) of subspace codes of maximum size when the remaining parameters are fixed. In contrast with its classical counterpart, however, much less is known

1991 *Mathematics Subject Classification*. Primary: 94B05, 05B25, 51E20; Secondary: 51E14, 51E22, 51E23.

Key words and phrases. Galois geometry, subspace code, linear operator channel, network coding, Gabidulin code, echelon-Ferrers construction, expurgation-augmentation, subspace polynomial, Dickson invariants.

Research supported by the National Natural Science Foundation of China under Grants 60872063 and 61571006.

¹In the special case $V = \mathbb{F}_q^v$ one also refers to V , v , q as *packet space*, *packet length* and *packet alphabet size*, respectively. This is motivated by the applications in network coding.

about the Main Problem of Subspace Coding, in particular in the general mixed-dimension case. For the current state of knowledge we refer to the online tables at subspacecodes.uni-bayreuth.de, a recently established service of the research group at Bayreuth University; cf. also [27]. For the mixed-dimension case, see [30] and the references therein.

In this paper we restrict ourselves to (a subcase of) the constant-dimension case of the Main Problem, which is somewhat more accessible and, because of its geometric significance, has been investigated earlier by researchers in Finite Geometry and solved in special cases. A subspace code in V is said to be a *constant-dimension code* if all its members have the same dimension k . For the parameters of a constant-dimension code \mathcal{C} we use the notation $(v, M, d; k)$ or $(v, M, d; k)_q$, where q, v, k have the same meaning as above, $M = \#\mathcal{C} = |\mathcal{C}|$, and d denotes the minimum (subspace) distance of \mathcal{C} . The Main Problem, restricted to the constant-dimension case, is to determine the maximum sizes $A_q(v, d; k)$ of $(v, M, d; k)_q$ codes.

For $X, Y \in \begin{bmatrix} V \\ k \end{bmatrix}$, the set of k -dimensional subspaces (“ k -subspaces”) of V , the formula for the subspace distance reduces to $d_s(X, Y) = 2k - 2\dim(X \cap Y) \in 2\mathbb{Z}$, and the inequality $d_s(X, Y) \geq d = 2\delta$ is equivalent to $\dim(X \cap Y) < t$ with $t = k - \delta + 1$. Hence a $(v, M, d; k)_q$ constant-dimension code \mathcal{C} with ambient vector space V may also be viewed as a set of $(k - 1)$ -flats in $\text{PG}(V)$ with $\#\mathcal{C} = M$ and the following property: $t = k - \delta + 1$ is the smallest integer such that every $(t - 1)$ -flat of $\text{PG}(V)$ (t -subspace of V) is contained in at most one member of \mathcal{C} . Finding a $(v, M, 2\delta; k)_q$ code of maximum size thus translates exactly into the packing problem for the incidence structure of t -subspaces versus k -subspaces of V , provided we identify “blocks” of this incidence structure (i.e. subspaces $X \in \begin{bmatrix} V \\ k \end{bmatrix}$) with sets of “points” (the set of all $T \in \begin{bmatrix} V \\ t \end{bmatrix}$ satisfying $T \subseteq X$).

The special case $t = 1$ of the restricted Main Problem asks for the maximum number of k -subspaces of V that are pairwise disjoint as point sets. In Finite Geometry such sets of subspaces are known as *partial spreads* and have been the subject of extensive research since the fundamental work of Beutelspacher [7]; see [18] for a survey. However, apart from the “spread case” $v \equiv 0 \pmod{k}$ and the line case $k = 2$ already solved in [7], the numbers $A_q(v, 2k; k)$ remain unknown in general. Only for certain specific parameter combinations they have been determined exactly—in the case $v \equiv 1 \pmod{k}$ [7], the case $q = 2, k = 3$ [19], and recently in the case $q = 2, v \equiv 2 \pmod{k}$ [35].

Predating the work of Koetter-Kschischang [32] on network coding, the Main Problem for general constant-dimension subspace codes has already been formulated and investigated by Metsch in the language of Finite Geometry (see [37], in particular Section 4) and by Wang, Xing and Safavi-Naini in their work on linear authentication codes (see [46, Th. 4.1] and the discussion following this theorem). The publication of [32] and the related work [41, 42] sparked a lot of research interest in constant-dimension subspace coding, focusing either on the derivation of upper bounds for the numbers $A_q(v, 2\delta; k)$ or on explicit code constructions, which provide lower bounds for $A_q(v, 2\delta; k)$. An non-authoritative, non-exhaustive selection of additional references is [33, 48, 24, 20, 14, 44].

The exact determination of $A_q(v, 2\delta; k)$ seems to be a very difficult problem even for moderate parameter sizes. To the best of our knowledge, there are currently only two parameter sets $(v, 2\delta; k)_q$ with $1 < t = k - \delta + 1 < k$, where $A_q(v, 2\delta; k)$ is known exactly: $A_2(6, 4; 3) = 77$ [31] and $A_2(13, 4; 3) = 1,597,245$ [9]. The

$(13, 1597245, 4; 3)_2$ code—in fact there exist many non-isomorphic codes with these parameters—is particularly remarkable, since it forms the first nontrivial example of a Steiner system over a finite field (a 2-analogue of the projective plane of order 3).

Our contribution to the restricted Main Problem in this paper pertains also to the case $q = 2$, $k = 3$, $d = 4$. In geometric terms, we consider sets of planes in a binary projective geometry $\text{PG}(V) \cong \text{PG}(v-1, \mathbb{F}_2)$ mutually intersecting in at most a point. We refer to these sets as (binary) *plane subspace codes*, and our interest is in finding the largest plane subspace code(s) over \mathbb{F}_2 with fixed packet length v . Before stating our main result, we shall briefly review previous work on binary plane subspace codes.

The exact results known in this case are the “trivial” (inasmuch as it reduces to a case with $t = 1$) $A_2(5, 4; 3) = A_2(5, 4; 2) = 9$, and the two already mentioned results $A_2(6, 4; 3) = 77$, $A_2(13, 4; 3) = 1,597,245$. In the smallest open case $v = 7$ we know $329 \leq A_2(7, 4; 3) \leq 381$. The lower bound stems from the computer-aided group-invariant construction in [11] (cf. [36, 29] for alternative constructions), and the upper bound is the size of a putative 2-analogue of the projective plane of order 2, whose existence is still undecided (despite the known solution in the much larger case $v = 13$). In addition, for $7 < v < 13$ strong lower bounds for $A_2(v, 4; 3)$ are known from the computational work in [10], which employs dedicated combinatorial optimization techniques for group-invariant subspace codes.

The best known constructive² lower bound for general v is provided by the echelon-Ferrers construction and its variants [21, 45, 22, 40]. It asserts that

$$\begin{aligned} A_2(v, 4; 3) &\geq 2^{2(v-3)} + \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2 = 2^{2v-6} + \frac{(2^{v-3} - 1)(2^{v-4} - 1)}{3} \\ &= 2^{2v-6} + 2^{2v-9} + 2^{2v-11} + \dots + 2^1 + 1 - 2^{v-4} \end{aligned}$$

for $v \leq 11$, with a slightly inferior bound for larger v . The quantity $2^{2(v-3)} + \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2$ also provides an upper bound for any $(v, M, 4; 3)_2$ code that contains a lifted maximum rank distance code (LMRD code) as a subcode and hence represents essentially the optimum achievable by the echelon-Ferrers construction and its variants [22]. Subsequently we will refer to this upper bound as the *LMRD code bound*.

The best known upper bound for unrestricted $(v, M, 4; 3)_2$ codes, a consequence of the known maximum size of partial line spreads in $\text{PG}(v-1, \mathbb{F}_2)$, is

$$(1) \quad A_2(v, 4; 3) \leq \begin{cases} \left\lfloor \frac{(2^v-1)(2^{v-1}-1)}{21} \right\rfloor & \text{for } v \equiv 1 \pmod{2}, \\ \left\lfloor \frac{(2^v-1)(2^{v-1}-5)}{21} \right\rfloor & \text{for } v \equiv 0 \pmod{2}, \end{cases}$$

and is substantially larger.³ It can be verified that the base-2 representation of the upper bound has the form $2^{2v-6} + 2^{2v-7} + 2^{2v-12} + 2^{2v-13} + \dots + 2^{2v-6s} + 2^{2v-6s-1} + \text{smaller terms}$, where $s = \lfloor (v-3)/6 \rfloor$.⁴

²For large v a non-constructive lower bound, which asymptotically matches the upper bound stated in the next paragraph, has been shown in [8].

³The maximum size of partial line spreads in $\text{PG}(v-1, \mathbb{F}_q)$ is known for all prime powers $q > 1$ [7], leading to an analogous (best known) upper bound for $A_q(v, 4; 3)$.

⁴We can also express these bounds asymptotically for $v \rightarrow \infty$ as $2^{2v-6} (\frac{7}{6} + o(1)) \leq A_2(v, 4; 3) \leq 2^{2v-6} (\frac{32}{21} + o(1))$.

Our main theorem, stated below, improves upon the echelon-Ferrers construction and all its variants for $q = 2$, $k = 3$ and infinitely many packet lengths v . The codes are constructed using a generalization of the expurgation-augmentation method introduced in [31, 36, 29]. Moreover, the augmentation step can be done in such a way that the resulting codes are invariant under a $(v - 3)$ -dimensional Singer subgroup of $\text{GL}(v, 2)$ acting trivially on the complementary three coordinates. More precisely, the codes have ambient space $V = W \times \mathbb{F}_{2^n}$, where $n = v - 3$ and W is a certain 3-dimensional \mathbb{F}_2 -subspace of \mathbb{F}_{2^n} , and are invariant under the group $\Sigma_v \leq \text{GL}(V)$ consisting of all maps of the form $(x, y) \mapsto (x, ry)$ with $r \in \mathbb{F}_{2^n}^\times$.

Main Theorem. (i) For $v \equiv 7 \pmod{8}$, there exists a Σ_v -invariant $(v, M, 4; 3)_2$ subspace code with

$$M \geq 2^{2(v-3)} + \frac{9}{8} \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2,$$

and consequently we have $A_2(v, 4; 3) \geq 2^{2(v-3)} + \frac{9}{8} \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2$ in this case.

(ii) For $v \equiv 3 \pmod{8}$, $v \geq 11$, there exists a Σ_v -invariant $(v, M, 4; 3)_2$ subspace code with

$$M \geq 2^{2(v-3)} + \frac{81}{64} \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2,$$

and consequently we have $A_2(v, 4; 3) \geq 2^{2(v-3)} + \frac{81}{64} \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2$ in this case.

The route to our main theorem is long and involved, at least when following the chronological order in which the various pieces were put together. We have deliberately left this order untouched, since it captures best the line of argument and the motivation for each subsequent step. In order to make the paper essentially self-contained, we provide an exposition of the basic and refined expurgation-augmentation method as developed in [31, 36, 29], including the key examples for $q = 2$. The theoretical analysis is supplemented⁵ by extensive computations, which were done using the computer algebra system SageMath (www.sagemath.org).⁶ In some cases the largest subspace codes obtained during this optimization process considerably exceed the bounds stated in the main theorem, and for packet lengths $v > 13$ they probably form the largest subspace codes explicitly known at present. For details we refer to Table 1.

In the remainder of this introduction we will provide a brief overview of the route to the main theorem and at the same time explain how the rest of this paper is organized.

The background of the expurgation-augmentation method, its previous developments, and the adaption to packet lengths $v > 7$ (the initial stage of our research) is described in Sections 2, 3, and 4. A key ingredient, dating back to [31], is a particular choice of the ambient space $V \cong \mathbb{F}_2^v$, which takes the structure of the optimal $(6, 77, 4; 3)_2$ subspace codes into account and, similar to the Singer representation, allows for using the multiplicative structure of a large finite field: We always take

⁵In fact the main result would have never been discovered without the computational data, which suggested the route pursued in later sections, and the need to make our computation more efficient.

⁶SageMath has proved to be an extremely versatile tool in our present research.

V as $W \times \mathbb{F}_{2^n}$, where $n = v - 3$ and W is some 3-subspace of \mathbb{F}_{2^n} , i.e., a plane in $\text{PG}(\mathbb{F}_{2^n})$.⁷

At the beginning, with a rough goal of generalizing the results of [29], we made computational experiments in the case $v = 8$; these were all but encouraging.⁸ But in fact, $v = 8$ and $v = 10$ are the only cases within the range $v \in \{7, 8, \dots, 16\}$ now covered by the computational part of our work, in which the refined (“rotation-invariant”) expurgation-augmentation method is inferior to the echelon-Ferrers construction and its variants. Our experiments used results in [29, Sect. 5], which express the number of planes that can be “locally” added to the expurgated lifted Gabidulin code without decreasing the minimum distance as the number of distinct values of a certain numerical invariant for planes in $\text{PG}(\mathbb{F}_{2^n})$, named σ -invariant in [29].

Although the σ -invariant and some of its properties generalize to arbitrary v , the case $v \geq 8$ differs fundamentally from the case $v = 7$ considered in [29] in the following respects:

- For $v \geq 8$ the Gabidulin code and the σ -invariant generally depend on the plane W used in $V = W \times \mathbb{F}_{2^n}$. As a consequence, all planes W in $\text{PG}(\mathbb{F}_{2^n})$ have to be taken into account for the subspace code optimization and there will certainly be no analogue of the nice explicit formula for the σ -invariant established for $v = 7$ in [29, Lemma 7].⁹
- For $v \geq 8$ there is no longer a canonical choice for the expurgated Gabidulin code in the refined expurgation-augmentation method. Instead there are $2^{v-6} - 1$ minimal subsets of \mathcal{G}_W , any combination of which can be removed to obtain an expurgated Gabidulin code. The subspace code optimization algorithm should consider all such combinations and select the best one.

With these two guidelines at hand, we created a simple prototype of the subspace code optimization algorithm, which generated planes W randomly, evaluated the associated invariant σ_W on all planes intersecting W in a line,¹⁰ and computed an optimal solution of the resulting optimization problem—maximize the difference between the image size of σ_W and the (suitably normalized¹¹) number of planes removed from \mathcal{L}_W , subsequently referred to as the *local net gain*—in a brute-force manner by exhaustive search through all combinations of minimal subsets of \mathcal{G}_W . With this algorithm and the additional observation that the problem setting is invariant under a fairly large group of collineations of $\text{PG}(\mathbb{F}_{2^n})$ acting on the set of planes W , we were able to solve the cases $v = 8, 9, 10$ completely and do a partial search for length $v = 11$. It turned out that the initial estimate, based on $v = 8$, had been too pessimistic; for $v = 9, 11$ our algorithm found solutions which exceeded the LMRD code bound.

⁷“ $\text{PG}(\mathbb{F}_{2^n})$ ” will be used as a shorthand for $\text{PG}(\mathbb{F}_{2^n}/\mathbb{F}_2)$, the projective geometry derived from the field extension $\mathbb{F}_{2^n}/\mathbb{F}_2$.

⁸For $v = 8$ the best subspace codes obtained by expurgation-augmentation have size $1024 + 93 = 1117$. Even if we were lucky to extend the codes by the theoretical maximum of $\begin{bmatrix} 5 \\ 2 \end{bmatrix}_2 = 155$ further planes, we would still remain way below the best known $(7, 1312, 4; 3)_2$ codes at that time. The currently best known $(7, M, 4; 3)_2$ has size $M = 1326$; cf. [10].

⁹Subsequently we will write \mathcal{G}_W , σ_W to indicate the dependence of the Gabidulin code, respectively, σ -invariant on W .

¹⁰The domain of σ_W consists precisely of those planes.

¹¹The total number of planes removed from \mathcal{L}_W is divided by the number of points on the special flat S defined in Section 2, in order to match the present “local” point-of-view.

The next step was to replace the exhaustive search through all $2^{2^{v-6}-1} - 1$ nonempty combinations of minimal subsets, which is clearly prohibitive for $v = 11$, by something more efficient. For this we inspected the data structure containing the computed values of σ_W , a matrix of size $(2^{n-3} - 1) \times (2^n - 1)$, $n = v - 3$, indexed with the solids (4-subspaces) T in $\text{PG}(\mathbb{F}_{2^n})$ containing W and the elements $a \in \mathbb{F}_{2^n}^\times$, which has as (i, j) -entry the number of planes E such that $E + W = T_i$ and $\sigma(E) = a_j$.¹² The matrix turned out to have a very special structure. Obviously it is divided into two parts representing elements $a \in \mathbb{F}_{2^n}^\times$ with ≤ 1 and > 1 preimages E under σ_W , respectively, and an optimal solution of the optimization problem must include all planes E of the first kind.¹³ Hence attention can be restricted to the second part, which turned out to be a square matrix of order $2^{n-3} - 1$ with columns indexed by an $(n - 3)$ -subspace (“collision space”) of \mathbb{F}_{2^n} . This submatrix and subspace will be called *collision matrix*, respectively, *collision space* of W or σ_W ; cf. Theorem 6.2 and Definition 6.3. As it turned out, collision matrices have only 3 different column shapes with nonzero entry patterns 4^1 , 2^3 , 1^7 and such that the supports of each column, viewed as a set of points in $\text{PG}(\mathbb{F}_{2^n}/W)$, forms a subspace (point, line, or plane) as well. The proof of these properties, which were first noticed through experiments, is given in Theorem 6.5.

The rather difficult proofs of the preceding (and also subsequent) observations use properties of so-called subspace polynomials and in particular the linear (i.e., degree one) coefficients of such polynomials, which are analogous to the elementary symmetric polynomials $\sigma_k(X) = X_1 X_2 \cdots X_k$. The basic theory of subspace polynomials and their coefficients, the “Dickson invariants”, has been developed long ago by L.E. DICKSON and O. ORE, cf. [17, 39]. We provide a brief account of this theory in Section 5, tailored to the case $q = 2$ and including some new (or at least less well-known) results, which are needed in subsequent sections.

Subspace polynomials have recently been used in other contexts [4, 3, 13], including a direct application to subspace coding [2]. We also offer a tiny but curious result in this direction (Corollary 3).

Armed with the new theoretical insight we were able to reformulate the local net gain maximization as a combinatorial optimization problem with a fairly rich structure provided by the collision matrix (Theorem 6.4, Corollary 4),¹⁴ and solve it completely for packet lengths $v \leq 13$. This and further non-exhaustive computations for $v \in \{14, 15, 16\}$ confirmed that the expurgation-augmentation method produces codes larger than the LMRD code bound for $v \neq 8, 10$; see Section 7.

The last (but not least) steps towards the main theorem were the following: Along with the maximum net gain computations we had recorded some statistical data for the collision matrices, with the goal of understanding which algebraic properties of W are responsible for a large maximum net gain. From this we noticed that the “best” planes were those with an entry “4” in their collision matrix, a property that can be described algebraically (Theorem 6.7), and with the largest number of so-called “missing points” (this concept is defined at the beginning of Section 6) in

¹²The solids correspond to the minimal subsets of \mathcal{G}_W that can be combined. Hence it is not necessary to distinguish between different planes in T_i at this stage.

¹³That is, there is no plane $E' \neq E$ intersecting W in a line and such that $\sigma(E) = \sigma(E')$.

¹⁴The problem bears some similarity to the set cover problems studied in Theoretical Computer Science; see [25].

their collision space. The best example for $v = 11$, in which W is equal to the trace-zero subspace of the subfield $\mathbb{F}_{16} \subset \mathbb{F}_{2^8}$ and whose 31×31 collision matrix is shown in Figure 1, helped us to understand that the geometric configuration formed by the multiset of missing points in the collision space determines the row-sum spectrum of the collision matrix and hence, via Corollary 4, to some extent controls the maximum achievable net gain. The precise relation is described in Theorem 6.6. Moreover, the trace-zero subspace of \mathbb{F}_{16} provides a natural candidate for W in all cases $v \equiv 3 \pmod{4}$, since for such v the field \mathbb{F}_{2^n} contains \mathbb{F}_{16} . The final steps where to show that the maximum net gain achievable with the trace-zero subspace satisfies the bounds stated in the main theorem. This is accomplished in Section 8, first in the case $v \equiv 7 \pmod{8}$, which is considerably easier, (Theorem 8.1) and then in the case $v \equiv 3 \pmod{8}$ (Theorem 8.2).

Theorem 6.6 encompasses the nice fact that the row-sum spectrum of a collision matrix can be computed in much the same way as the weight distribution of a linear code from geometric information about an associated multiset of points in some projective geometry. This connection, first described after Theorem 6.6 and used in the proofs of Theorems 8.1 and 8.2, is made more explicit in Section 9. Here we show, by bounding the maximum net gain in terms of a certain quantity (“code sum”) and estimating this code sum for all projective binary linear $[\mu, k]$ codes of length $\mu \leq 7$, that the non-exhaustively computed maximum net gains for $v = 14, 15$ in Table 1 represent the true maximum (Theorem 9.3).

Based on the accumulated computational data, we conjecture that the largest subspace codes obtained by the expurgation-augmentation method exceed the LMRD code bound for all sufficiently large packet lengths v (Conjecture 1). For odd v this conjecture is strongly supported by computational data on the distribution of missing points in the collision space; see the end of Section 8.

The paper concludes with Section 10, which provides a discussion of some in a sense “neglected” aspects of our work and gives some suggestions for future research.

Some familiarity with basic concepts and terminology from Finite Geometry is indispensable for understanding this paper. The relevant background information can be found in [15], [28], or [6]. Regarding notation, we only mention at this point the abbreviation $\text{Tr}(x) = \text{Tr}_{\mathbb{F}_{2^n}/\mathbb{F}_2}(x) = x + x^2 + x^4 + \cdots + x^{2^{n-1}}$ for the absolute trace of \mathbb{F}_{2^n} , which is used frequently in the sequel. All other non-standard notation will be explained on its first occurrence.

2. PRELIMINARIES ON PLANE SUBSPACE CODES

Throughout this section let \mathcal{C} be a plane subspace code with parameters $(v, M, 4; 3)$ and ambient space V , where w.l.o.g. $v \geq 6$. Since planes in \mathcal{C} do not have a line in common, \mathcal{C} covers (i.e., its members contain) precisely $7M$ lines of $\text{PG}(V)$. Conversely, if we know the number l of lines covered by \mathcal{C} , we can recover the size of \mathcal{C} as $M = l/7$. Hence maximizing M and l are equivalent problems.

This point of view is especially useful when looking at lifted maximum rank distance codes (LMRD codes) with these parameters. Such codes are obtained from maximum rank-distance-2 matrix codes in $\mathbb{F}_2^{3 \times (v-3)}$, e.g. Gabidulin codes, by the *lifting construction* $\mathbf{A} \mapsto \langle (\mathbf{I}_3 | \mathbf{A}) \rangle$ (“prepending the 3×3 identity matrix to \mathbf{A} and then taking the row space”) and have size $2^{2(v-3)}$. The planes obtained in this way are disjoint from the special $(v-3)$ -dimensional subspace $S = \{\mathbf{x} \in \mathbb{F}_2^v; x_1 = x_2 = x_3 = 0\}$ and, as remarked above, cover $7 \cdot 2^{2(v-3)}$ lines of $\text{PG}(\mathbb{F}_2^v)$. On the

other hand, standard counting facts in finite projective spaces imply that the total number of lines in $\text{PG}(\mathbb{F}_2^v)$ disjoint from S is also $7 \cdot 2^{2(v-3)}$. Hence any LMRD code in \mathbb{F}_2^v forms a perfect cover of the set of lines disjoint of S (and conversely, any such perfect cover arises from a maximum rank-distance-2 matrix code in the way described).

This leads directly to the LMRD code bound mentioned in Section 1: If \mathcal{C} contains an LMRD code then it cannot contain planes meeting S in a point (since these contain lines disjoint from S) and hence contains at most $\begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2$ further planes, one for each line contained in S .¹⁵

In order to overcome the LMRD code bound, we should therefore start with a smaller set of planes disjoint from S . It is reasonable to choose this set as a large subcode of a Gabidulin code, and it has been shown in [31, 36, 29] that this idea can indeed be put to work for $v = 6, 7$. The method, which we call *expurgation-augmentation*—remove some “old” planes from the Gabidulin code (“expurgate” the Gabidulin code) and add in turn some “new” planes meeting S in a point (“augment” the expurgated Gabidulin code)—, is described in the next section. Crucial for the success of the method is a particular choice of the ambient space V , which involves a large extension field of \mathbb{F}_2 and allows us later to employ properties of linearized polynomials and the multiplicative structure of the extension field. This choice of V will be discussed in the remainder of this section.

The ambient space for the expurgation-augmentation method is taken as $V = W \times \mathbb{F}_{2^n}$, $n = v - 3$, for some 3-dimensional \mathbb{F}_2 -subspace W of \mathbb{F}_{2^n} .¹⁶ In this model the special $(v - 3)$ -subspace is $S = \{0\} \times \mathbb{F}_{2^n}$, and W is represented within V as $\widetilde{W} = W \times \{0\}$.

The corresponding Gabidulin code can be defined in a basis-independent manner as $\mathcal{G}_W = \{a_0x + a_1x^2; a_0, a_1 \in \mathbb{F}_{2^n}\}$, where $a_0x + a_1x^2$ is used as an abbreviation for the \mathbb{F}_2 -linear map $W \rightarrow \mathbb{F}_{2^n}$, $x \mapsto a_0x + a_1x^2$. The lifted Gabidulin code in this model is $\mathcal{L}_W = \{\Gamma_f; f \in \mathcal{G}_W\}$, where $\Gamma_f = \{(x, f(x)); x \in W\}$ denotes the graph of f (in the sense of Real Analysis, if you like).

The $7 \cdot 2^{2(v-3)} = 7 \cdot 2^{2n}$ lines covered by \mathcal{L}_W are precisely the graphs of the restrictions $f|_Z: Z \rightarrow \mathbb{F}_{2^n}$, $x \mapsto f(x)$ of $f \in \mathcal{G}_W$ to lines (2-subspaces) $Z \subset W$. The perfect cover property of \mathcal{L}_W is reflected in the fact that the maps $\mathcal{G}_W \rightarrow \text{Hom}(Z, \mathbb{F}_{2^n})$, $f \mapsto f|_Z$ are linear isomorphisms,¹⁷ and hence any line disjoint from S , which is the graph of a unique \mathbb{F}_2 -linear map $g: Z \rightarrow \mathbb{F}_{2^n}$ for some Z , is covered precisely once by \mathcal{L}_W .

More generally, any \mathbb{F}_2 -subspace U of V can be parametrized in the form

$$(2) \quad U = \{(x, f(x) + y); x \in Z, y \in T, f \in \text{Hom}(Z, \mathbb{F}_{2^n})\},$$

where $Z \subseteq W$, $T \subseteq \mathbb{F}_{2^n}$ are \mathbb{F}_2 -subspaces and f is an \mathbb{F}_2 -linear map. We write $U = U(Z, T, f)$ in this case. The spaces Z , W are recovered from U as $Z = \{x \in W; \exists y \in \mathbb{F}_{2^n} \text{ such that } (x, y) \in U\}$ and $T = \{y \in \mathbb{F}_{2^n}; (0, y) \in U\}$. The map f can be any element of $\text{Hom}(Z, \mathbb{F}_{2^n})$ satisfying $\Gamma_f \subseteq U$ and corresponds

¹⁵Since a plane contained in S covers 7 lines, it is also clear that \mathcal{C} meets the bound with equality iff it has a subcode forming a perfect cover of the lines in S .

¹⁶Note that $v - 3 \leq n$ in view of our assumption $v \geq 6$.

¹⁷As usual, $\text{Hom}(X, Y)$ denotes the vector space of all linear maps from X to Y . Here the ground field is \mathbb{F}_2 , and $\text{Hom}(Z, \mathbb{F}_{2^n})$ is considered as a vector space over \mathbb{F}_2 (although it is also a vector space over \mathbb{F}_{2^n}).

to a complement for $U \cap S = \{0\} \times T$ in U via $f \mapsto \Gamma_f$.¹⁸ Further, we have $U(Z, T, f) = U(Z', T', f')$ if and only if $Z = Z'$, $T = T'$ and $f - f' \in \text{Hom}(Z, T)$ (i.e., $f(x) - f'(x) \in T$ for all $x \in Z$). The parametrization $U = U(Z, T, f)$ thus induces a 1-1 correspondence between \mathbb{F}_2 -subspaces of V and triples $(Z, T, f + \text{Hom}(Z, T))$. Finally, the incidence relation on subspaces of V translates into the following conditions for the parameters: $U(Z, T, f) \subseteq U(Z', T', f')$ iff $Z \subseteq Z'$, $T \subseteq T'$ and $f'|_Z - f \in \text{Hom}(Z, T')$.

3. THE BASICS OF EXPURGATION-AUGMENTATION

The underlying geometric idea is to find sets of planes in \mathcal{L}_W , whose lines can be rearranged into new planes meeting S in a point. Removing the planes in such a set from \mathcal{L}_W and adding in turn the new planes to the expurgated subspace code preserves the exact cover property with respect to lines disjoint from S . Moreover, if t planes are removed then $7t$ lines disjoint from S are involved and, since new planes contain only 4 such lines, the subspace code size increases by $\frac{7t}{4} - t = \frac{3t}{4}$.¹⁹ However, we must be careful to avoid any multiple cover of a line meeting S in a point.

Planes of $\text{PG}(V)$ meeting S in a point $P = \mathbb{F}_2(0, r)$ are parametrized as $N = U(Z, \mathbb{F}_2 r, g)$ for some line Z in W and some linear map $g: Z \rightarrow \mathbb{F}_2^n$. Using the natural isomorphism $S = \{0\} \times \mathbb{F}_2^n \cong \mathbb{F}_2^n$, we may view P as a point of $\text{PG}(\mathbb{F}_2^n)$ and write the parametrization as $N = U(Z, P, g)$. The line Z specifies the hyperplane $H = N \vee S = U(Z, \mathbb{F}_2^n, 0)$ above S that contains N , and the 4 lines in N disjoint from S are the graphs of the rank-distance-1 clique $g + \text{Hom}(Z, P)$. At this point it comes in handy that $f \mapsto f|_Z$ identifies \mathcal{G}_W with $\text{Hom}(Z, \mathbb{F}_2^n)$. The associated map on planes is $\Gamma_f \mapsto \Gamma_{f|_Z} = \Gamma_f \cap H$ and can be used to determine exactly the set of 4 planes in \mathcal{L}_W that determine the 4 lines Γ_h , $h \in g + \text{Hom}(Z, P)$.

Before stating the criterion for exact rearrangement, it will be convenient to introduce a few special \mathbb{F}_2 -subspaces of \mathcal{G}_W . We define

$$\begin{aligned} (3) \quad \mathcal{T} &= \{ux^2 + u^2x; u \in W\}, \\ (4) \quad \mathcal{R} &= \{ux^2 + u^2x; u \in \mathbb{F}_2^n\}, \\ (5) \quad \mathcal{D}(Z, P) &= \{f \in \mathcal{G}_W; f(Z) \subseteq P\}, \end{aligned}$$

the latter with Z, P having the same meaning as above. Since the nonzero maps in \mathcal{R} have the factorization $ux(x + u)$, we have $\text{Ker}(f) = \mathbb{F}_2 u$, $\text{rk}(f) = 2$ if $f \in \mathcal{T}$ and $\text{Ker}(f) = \{0\}$, $\text{rk}(f) = 3$ if $f \in \mathcal{R} \setminus \mathcal{T}$.²⁰ Further, since $\mathcal{D}(Z, P)$ is mapped to $\text{Hom}(Z, P)$ by $f \mapsto f|_Z$, it is clear that $\#\mathcal{D}(Z, P) = 4$. Writing $Z = \langle a, b \rangle = \{0, a, b, a+b\}$, it is easy to verify that $\mathcal{D}(Z, \mathbb{F}_2(ab^2 + a^2b)) = \{0, ax^2 + a^2x, bx^2 + b^2x, (a+b)x^2 + (a+b)^2x\}$ and, using the factorized form,

$$\mathcal{D}(Z, \mathbb{F}_2 r) = \left\{ 0, \frac{rax(x+a)}{ab(a+b)}, \frac{rbx(x+b)}{ab(a+b)}, \frac{r(a+b)x(x+a+b)}{ab(a+b)} \right\}$$

in general.

¹⁸An explicit map f is obtained by choosing a basis B of Z and defining $f(b)$ as any y such that $(b, y) \in U$.

¹⁹Let us keep in mind that for beating the LMRD code bound we should have $t \geq \frac{4}{3} \lfloor \frac{n}{2} \rfloor_2 = \frac{(2^{n+1}-1)(2^n-1)}{9} \approx \frac{2}{9} \#\mathcal{G}_W$.

²⁰In the original version in [31], which was designed for $v = 6$, the subspace W is equal to $\mathbb{F}_{2^n} = \mathbb{F}_8$ and \mathcal{R} coincides with \mathcal{T} .

The preceding considerations imply the following

Lemma 3.1 (cf. [31, Lemma 10] and [29, Lemma 4]). *Let $\mathcal{A} \subseteq \mathcal{G}_W$ be a subset of size t . The $7t$ lines contained in the members of $\{\Gamma_f; f \in \mathcal{A}\} \subseteq \mathcal{L}_W$ can be rearranged into $7t/4$ new planes meeting S in a point if and only if $t = 4m$ is a multiple of 4 and for every line $Z \subset W$ there exist (not necessarily distinct) points $P_1, \dots, P_m \in S$ and linear maps $f_1, \dots, f_m \in \mathcal{A}$ such that*

$$\mathcal{A} = \biguplus_{i=1}^m (f_i + \mathcal{D}(Z, P_i)).$$

In other words, \mathcal{A} should admit decompositions into disjoint cosets of spaces $\mathcal{D}(Z, \cdot)$ simultaneously for each Z . The points P_1, \dots, P_m may coincide, in which case the condition reduces to a representation of \mathcal{A} as a union of cosets of some space $\mathcal{D}(Z, P)$.

The criterion in Lemma 3.1 seems to be rather complicated to check and a description of all such rearrangements for any given subset \mathcal{A} out of reach. However, there is an obvious candidate for \mathcal{A} that admits a simultaneous decomposition of the required form, viz. the space \mathcal{T} , which contains the 7 subspaces $\mathcal{D}(Z, \mathbb{F}_2(ab^2 + a^2b))$ and hence decomposes into 2 cosets of each of them. The corresponding rearrangement is clearly unique, and by using \mathcal{T} as the basic building block we obtain a large number of sets \mathcal{A} satisfying the condition in Lemma 3.1 for a specific decomposition. This will be sufficient for our purposes.

Every binomial $a_0x + a_1x^2 \in \mathcal{G}_W$ is uniquely represented as $r(ux^2 + u^2x)$ with $r, u \in \mathbb{F}_{2^n}^\times$ (i.e., as rf with $r \in \mathbb{F}_{2^n}^\times$, $f \in \mathcal{R} \setminus \{0\}$).²¹ Hence \mathcal{G}_W consists of 0, the $2(2^n - 1)$ monomials rx, rx^2 , which have rank 3, the $7(2^n - 1)$ rank-2 binomials rf with $f \in \mathcal{T} \setminus \{0\}$, and the $(2^n - 8)(2^n - 1)$ rank-3 binomials rf with $f \in \mathcal{R} \setminus \mathcal{T}$. The subset of rank-3 binomials decomposes into $(2^{n-3} - 1)(2^n - 1)$ pairwise disjoint “rotated” cosets $r(f + \mathcal{T})$, where $r \in \mathbb{F}_{2^n}^\times$ and $f \in \mathcal{R} \setminus \mathcal{T}$ is determined modulo \mathcal{T} .

Just like \mathcal{T} , the set $r(f + \mathcal{T})$ admits a unique simultaneous decomposition into 2 cosets of $\mathcal{D}(Z, \mathbb{F}_2r(ab^2 + a^2b))$. Extending this in the obvious way to unions of rotated cosets, we see that any such union admits a simultaneous decomposition into cosets of spaces $\mathcal{D}(Z, \cdot)$, as required in Lemma 3.1. Hence we have the following

Lemma 3.2. *Suppose $\mathcal{A} \subseteq \mathcal{G}_W$ is the union of some of the $(2^{n-3} - 1)(2^n - 1)$ rotated cosets $r(f + \mathcal{T})$ ($r \in \mathbb{F}_{2^n}^\times$, $f \in \mathcal{R} \setminus \mathcal{T}$) and at most one rotated subspace $r\mathcal{T}$. Then the lines in $\{\Gamma_f; f \in \mathcal{A}\}$ can be exactly rearranged into new planes meeting S in a point.*

The corresponding “obvious” exact rearrangement of free lines into new planes will be called the *standard rearrangement*.²²

The previous construction provides us with myriads of subsets $\mathcal{A} \subseteq \mathcal{G}_W$ satisfying the conditions in Lemma 3.1, but it does not tell us whether the set \mathcal{N} of new planes of the corresponding standard rearrangement has $d_s(\mathcal{N}) \geq 4$. We are interested in the largest subsets \mathcal{A} having this extra property, since the size $\#\mathcal{C} = 4^{v-3} + 3t/4$ of the modified subspace code

$$\mathcal{C} = \mathcal{L}_W - \{\Gamma_f; f \in \mathcal{A}\} \cup \mathcal{N},$$

²¹Solving $a_1x^2 + a_0x = ru^2x + ru^2x$ for R, u gives $u = a_0/a_1$, $r = a_1^2/a_0$.

²²“Free line” refers to a line covered by $\{\Gamma_f; f \in \mathcal{A}\}$. After removal of this set of planes from \mathcal{G}_W , such a line is uncovered, i.e., “free”. This name was coined in [31].

which then has $d_s(\mathcal{C}) \geq 4$ as well, is an increasing function of $t = \#\mathcal{A}$.

Rearrangement Problem (RP). Determine the subsets $\mathcal{A} \subseteq \mathcal{G}_W$ of maximum size that are unions of pairwise disjoint rotated cosets $r(f + \mathcal{T})$ (as in Lemma 3.2) and whose standard rearrangement into new planes forms a subspace code \mathcal{N} with $d_s(\mathcal{N}) \geq 4$.

We are not able to solve the rearrangement problem, but we will exhibit fairly large subsets \mathcal{A} with this property (cf. Theorem 3.4 below). The resulting modified subspace codes, however, are still inferior to those produced by the echelon-Ferrers construction (although it is conceivable that they can be extended by $\approx \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2$ further planes meeting S in a line to a code exceeding the LMRD code bound). That notwithstanding, the preparations made en-route to Theorem 3.4 will be needed for the refined approach taken up in Section 4.

Before proceeding, it will be convenient to discuss some properties of the map $\delta: \mathbb{F}_{2^n} \times \mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^n}$, $(x, y) \mapsto xy^2 + x^2y = xy(x+y)$. The map δ is \mathbb{F}_2 -bilinear, symmetric, and alternating (i.e., $\delta(x, x) = 0$ for $x \in \mathbb{F}_{2^n}$). Fixing the second argument, say, we have that $x \mapsto \delta(x, y)$, $y \neq 0$, is \mathbb{F}_2 -linear with kernel \mathbb{F}_2y and hence induces a collineation from the quotient geometry $\text{PG}(\mathbb{F}_{2^n})/P$, $P = \mathbb{F}_2y$, onto some hyperplane H in $\text{PG}(\mathbb{F}_{2^n})$.²³ Using $xy^2 + x^2y = y^3(x/y + (x/y)^2)$ and Hilbert's Satz 90, this hyperplane is easily seen to have equation $\text{Tr}(x/y^3) = 0$. The factorized form $\delta(x, y) = xy(x+y)$ reveals that $\delta(x, y)$ is equal to the product of the three (nonzero) points on the line $L = \langle x, y \rangle = \{0, x, y, x+y\}$ and thus provides a second geometric interpretation of $\delta(x, y)$. In particular, by setting $\delta(L) = xy(x+y)$ we obtain a map from lines to points of $\text{PG}(\mathbb{F}_{2^n})$. The following property of this map turns out to be crucial for the subsequent development; cf. Section 5.

Lemma 3.3. *$L \mapsto \delta(L)$ maps the lines contained in any plane E of $\text{PG}(\mathbb{F}_{2^n})$ bijectively onto the points of another plane E' . Moreover, the induced map from $\text{PG}(E)$ to $\text{PG}(E')$ is a correlation (i.e., an incidence reversing bijection mapping lines to points and points to lines).*

Proof. Writing $E = \langle a, b, c \rangle$, we must show that $E' = \langle ab^2 + a^2b, ac^2 + a^2c, bc^2 + b^2c \rangle = \langle \delta(a, b), \delta(a, c), \delta(b, c) \rangle$ has the required property. The lines of E are $L_1 = \overline{a, b}$, $L_2 = \overline{a, c}$, $L_3 = \overline{b, c}$, $L_4 = \overline{a, b+c}$, $L_5 = \overline{b, a+c}$, $L_6 = \overline{c, a+b}$, and $L_7 = \overline{a+b, a+c}$. Using the stated properties of $(x, y) \mapsto \delta(x, y)$, we obtain $\delta(L_4) = \delta(a, b) + \delta(a, c)$, $\delta(L_5) = \delta(a, b) + \delta(b, c)$, $\delta(L_6) = \delta(a, c) + \delta(b, c)$, $\delta(L_7) = \delta(a, b) + \delta(a, c) + \delta(b, c)$. Together with $\delta(L_i) \neq 0$ for $1 \leq i \leq 7$ this shows that E' is indeed a plane (i.e., $\delta(a, b)$, $\delta(a, c)$, $\delta(b, c)$ are linearly independent) and contains precisely the points $\delta(L_i)$, $1 \leq i \leq 7$, as claimed. Finally, $L \mapsto \delta(L)$ maps the three lines in E through a fixed point P onto some line in E' (since it induces a collineation $\text{PG}(\mathbb{F}_{2^n})/P \rightarrow H$), proving the last assertion. \square

Remark 1. For any line $L = \langle a, b \rangle$ in $\text{PG}(\mathbb{F}_{2^n})$ we may form the *line polynomial* $s_L(X) = \prod_{u \in L} (X - u) = X(X+a)(X+b)(X+a+b) = (X^2 + aX)(X^2 + aX + b(a+b)) = X^4 + (a^2 + ab + b^2)X^2 + ab(a+b)X \in \mathbb{F}_{2^n}[X]$. The coefficients of $s_L(X)$, viewed as polynomials in $\mathbb{F}_2[a, b]$, are $\text{GL}(2, \mathbb{F}_2)$ -invariants and freely generate the invariant ring $R = \mathbb{F}_2[a, b]^{\text{GL}(2, \mathbb{F}_2)}$ in the sense that $R = \mathbb{F}_2[a^2 + ab + b^2, ab(a+b)]$

²³The points and lines of $\text{PG}(\mathbb{F}_{2^n})/P \cong \text{PG}(\mathbb{F}_{2^n}/\mathbb{F}_2y)$ are the lines and planes of $\text{PG}(\mathbb{F}_{2^n})$ through P , respectively, and the incidence relation is the induced one.

is a polynomial ring. An analogous result holds for arbitrary subspaces (in place of lines) and prime powers $q > 1$ (in place of 2). This q -analogue of the fundamental theorem for symmetric polynomials is due to Dickson [17], and the coefficients $\delta_i^{(k)}$ of the generic k -dimensional subspace polynomial $s_U(X) = X^{q^k} - \delta_1^{(k)} X^{q^{k-1}} \pm \cdots + (-1)^k \delta_k^{(k)} X$ are accordingly referred to as q -ary, k -dimensional *Dickson invariants*. Thus $\delta(a, b) = ab^2 + a^2b = ab(a + b)$ is equal to the second binary 2-dimensional Dickson invariant (“line invariant”) $\delta_2^{(2)}$.

Now we resume our analysis of the rearrangement problem. Since $f \in \mathcal{R}$ has the form $f(x) = ux^2 + u^2x = \delta(u, x)$ and $f \in \mathcal{T}$ iff $u \in W$, we can write $r(f + \mathcal{T})$ as $r(\delta(u, x) + \delta(W, x)) = r\delta(u + W, x)$. The 14 new planes obtained from $r\delta(u + W, x)$ by the standard rearrangement are $U(Z, r\delta(Z), r\delta(u, x))$ and $U(Z, r\delta(Z), r\delta(u + c, x))$ with Z varying over the 7 lines in W and $c \in W \setminus Z$ (thus c depends on Z). Two distinct new planes (not necessarily from the same rotated coset) have a line in common if and only if they pass through the same point in S and have another point outside S in common. The 12 points outside S covered by $U(Z, r\delta(Z), r\delta(u, x))$ and $U(Z, r\delta(Z), r\delta(u + c, x))$, respectively, are

$$(6) \quad \begin{array}{ll} (a, r\delta(u, a)) & (a, r\delta(u + c, a)) \\ (a, r\delta(u + b, a)) & (a, r\delta(u + c + b, a)) \\ (b, r\delta(u, b)) & (b, r\delta(u + c, b)) \\ (b, r\delta(u + a, b)) & (b, r\delta(u + c + a, b)) \\ (a + b, r\delta(u, a + b)) & (a + b, r\delta(u + c, a + b)) \\ (a + b, r\delta(u + a, a + b)) & (a + b, r\delta(u + c + a, a + b)) \end{array}$$

(listed column-wise and writing $Z = \langle a, b \rangle$ as before).

Theorem 3.4. (i) The standard rearrangement of the 2^n , $n = v - 3$, planes in $\{\Gamma_f; f \in \mathcal{R}\}$ forms a set \mathcal{N} of new planes satisfying $d_s(\mathcal{N}) \geq 4$.
(ii) If $r_1, \dots, r_s \in \mathbb{F}_{2^n}^\times$ are such that $r_i W' \cap r_j W' = \emptyset$ for $1 \leq i < j \leq s$, then the standard rearrangement of the $2^n + (s - 1)(2^n - 8)$ planes Γ_f , $f \in r_1 \mathcal{R} \uplus r_2(\mathcal{R} \setminus \mathcal{T}) \uplus \cdots \uplus r_s(\mathcal{R} \setminus \mathcal{T})$ satisfies $d_s(\mathcal{N}) \geq 4$.
(iii) If $v \equiv 0 \pmod{3}$ and W is chosen as the subfield $\mathbb{F}_8 \subset \mathbb{F}_{2^n}$, then the standard rearrangement of the $2^n + (2^n - 8)^2/7$ planes Γ_f , $f \in \mathcal{R} \uplus r(\mathcal{R} \setminus \mathcal{T}) \uplus \cdots \uplus r^{(2^n-1)/7-1}(\mathcal{R} \setminus \mathcal{T})$ satisfies $d_s(\mathcal{N}) \geq 4$ and yields a modified subspace code \mathcal{C} with parameters $(v, 4^{v-3} + \frac{3}{7}(4^{v-4} - 9 \cdot 2^{v-5} + 16), 4; 3)_2$.

Proof. (i) Consider the 4 points in (6) with first coordinate a . If u varies over a set of coset representatives for \mathbb{F}_{2^n}/W , then the second coordinate of the 4 points takes precisely the values $\delta(x, a)$ with x varying over a set of coset representatives for $\mathbb{F}_{2^n}/\mathbb{F}_2 a$. Since δ is one-to-one on $\mathbb{F}_{2^n}/\mathbb{F}_2 a$, these 2^{n-1} values, and hence also the 2^{n-1} points $(a, \delta(x, a))$, are distinct. This reasoning applies to the points with first coordinate b or $a + b$ as well and shows that the 2^{n-2} planes in \mathcal{N} with the same Z pairwise intersect only in the point $(0, \delta(Z)) \in S$. But for $Z \neq Z'$ we have $(0, \delta(Z)) \neq (0, \delta(Z'))$ by Lemma 3.3, and hence planes in \mathcal{N} with different Z have subspace distance ≥ 4 as well. This completes the proof of (i).

(ii) The new planes in the standard rearrangement of $\{\Gamma_f; f \in r_i \mathcal{R}\}$ intersect S in the points of the plane $\{0\} \times r_i W'$. For $i \neq j$, since $r_i W' \cap r_j W' = \emptyset$ by

assumption, new planes obtained from $r_i\mathcal{R}$ and $r_j\mathcal{R}$ cannot intersect in S and hence have subspace distance ≥ 4 . Together with (i) this proves (ii).

(iii) For $W = \mathbb{F}_8$ we have $W' = W$, and the planes $W, rW, \dots, r^{(2^n-1)/7-1}W$ are pairwise disjoint.²⁴ Hence the first assertion in (iii) follows from (ii). Finally, the number of planes in \mathcal{C} is $4^n + \frac{3}{4}(2^n + \frac{1}{7}(2^n - 8)^2) = 4^n + \frac{3}{4 \cdot 7}(4^n - 9 \cdot 2^n + 64) = 4^n + \frac{3}{7}(4^{n-1} - 9 \cdot 2^{n-2} + 16)$, as claimed. \square

Remark 2. The following more geometric view of Theorem 3.4(i), which yields an alternative proof, may be of interest.

The “removed” set of planes $\{\Gamma_f; f \in \mathcal{R}\}$ covers precisely half of the points of $\text{PG}(V)$ outside S and forms an exact 2-cover of this set of points. This follows from the fact that $ux^2 + u^2x = y$, viewed as an equation for $u \in \mathbb{F}_{2^n}$ with parameter $x \neq 0$, has 2 solutions if $\text{Tr}(x^{-3}y) = 0$ and no solution if $\text{Tr}(x^{-3}y) = 1$. In other words, in each $(n+1)$ -dimensional space $F \supset S$, whose affine part is of the form $F \setminus S = \{x\} \times S$ for some nonzero $x \in W$, exactly the points $\mathbb{F}_2(x, x^3v)$ with $\text{Tr}(v) = 0$ are covered. There are 2^{n-1} such points, forming the affine part of an n -subspace of F intersecting S in $\{x^3v; \text{Tr}(v) = 0\}$.²⁵

Since $\text{rk}(f_1 - f_2) = 2$ iff $f_1 - f_2 \in \mathcal{T}$, this 2-cover is made up from smaller pieces $\{\Gamma_g; g \in f + \mathcal{T}\}$ corresponding to the cosets in \mathcal{R}/\mathcal{T} . The 8 planes in such a set mutually intersect in a point and hence 2-cover a set of $\binom{8}{2} = 28$ points (4 points in each F).

A point Q covered by $\{\Gamma_f; f \in \mathcal{R}\}$ is the intersection point of unique planes E_1 and E_2 in $\{\Gamma_f; f \in \mathcal{R}\}$. Each of E_1, E_2 contains 3 lines through Q , which represent the 3 hyperplanes above $F = Q \vee S$. Hence these lines are matched into 3 pairs; the lines in a pair determine the same hyperplane and generate a new plane in the standard rearrangement. It follows that the set \mathcal{N} of new planes of the standard rearrangement of $\{\Gamma_f; f \in \mathcal{R}\}$ forms a 3-cover of the same set of points and that new planes N_1, N_2 meeting in S do not meet outside S , since $N_1 \cap S = N_2 \cap S$ implies $N_1 \vee S = N_2 \vee S$, which in turn follows from the injectivity $Z = \langle a, b \rangle \mapsto P = \mathbb{F}_2(0, ab^2 + a^2b)$. This gives again Part (i) of Theorem 3.4.

Example 1 ($v = 6$). This is the smallest case, where Theorem 3.4 applies. Here $W = \mathbb{F}_8$, $\mathcal{R} = \mathcal{T}$, and Part (i) yields a subspace code $\mathcal{C} = \mathcal{L} \setminus \{\Gamma_f; f \in \mathcal{T}\} \cup \mathcal{N}$ of size $\#\mathcal{C} = 70$ consisting of the planes

$$G(a_0, a_1) = \{(x, a_0x + a_1x^2); x \in \mathbb{F}_8\}, \quad a_0, a_1 \in \mathbb{F}_8, \quad a_0 \neq a_1^2;$$

$$N(Z, c) = \{(x, cx^2 + c^2x + y\delta(Z)); y \in Z, \eta \in \mathbb{F}_2\}, \quad Z \subset \mathbb{F}_8 \text{ a line}, \quad c \in \mathbb{F}_8/Z.$$

This provides the essential step in the construction of an optimal $(6, 77, 4; 3)_2$ code of Type A in [31].

The construction is completed in the following way: The 28 points $\mathbb{F}_2(x, y)$ covered by the new planes $N(Z, c)$ outside $S = \{0\} \times \mathbb{F}_8$ are those covered by $G(a^2, a)$, $a \in \mathbb{F}_8$, and satisfy $\text{Tr}(x^4y) = 0$, as is clear from $x^4(ax^2 + a^2x) = x^{-3}(ax^2 + a^2x) = a/x + (a^2/x^2)$.²⁶ Now it is possible to connect the 7 lines in S to 7 points outside S

²⁴In fact they form the standard example of a plane spread in $\text{PG}(\mathbb{F}_{2^n})$.

²⁵Here we use again the identification $S = \{0\} \times \mathbb{F}_{2^n} \cong \mathbb{F}_{2^n}$.

²⁶The hyperbolic quadric \mathcal{H} in $\text{PG}(V) \cong \text{PG}(5, \mathbb{F}_2)$ with equation $\text{Tr}(x^4y) = 0$ consists of these 28 points and the 7 points in S . Together with S the 14 planes $N(Z, c)$ form one of the two sets of generators of \mathcal{H} . This provides the link to the alternative construction of a $(6, 77, 4; 3)_2$ code of Type A in [14].

in such a way that the resulting planes cover precisely the 28 points $\mathbb{F}_2(x, y)$ outside S satisfying $\text{Tr}(x^4y) = 1$. For this simply connect the point $\mathbb{F}_2(x, x^3)$, which has $\text{Tr}(x^4x^3) = \text{Tr}(1) = 1$, to the line $\{y \in \mathbb{F}_8; \text{Tr}(x^4y) = 0\}$. The resulting 7 planes can be added to \mathcal{C} to form the desired $(6, 77, 4; 3)_2$ code.

Theorem 3.4(iii) is still too weak to produce codes meeting (let alone exceeding) the LMRD code bound. More generally, any choice of \mathcal{A} that avoids new planes with different Z meeting in a point of S is subject to the bound $\#\mathcal{A} \leq \frac{1}{7}\#\mathcal{G}_W = \frac{1}{7}4^{v-3}$ and yields a net gain relative to $\#\mathcal{G}_W = 4^{v-3}$ of at most $\frac{3}{7}4^{v-4} < \binom{v-3}{2}_2 \approx \frac{2}{3}4^{v-4}$ planes. This remains true even if we relax the condition of exact rearrangement and augment the expurgated Gabidulin code by the maximum number of planes meeting S in a point while maintaining subspace distance ≥ 4 .

4. THE REFINED APPROACH

In this section we relax the condition of exact rearrangement but restrict attention to rotation-invariant subsets $\mathcal{A} \subseteq \mathcal{G}_W$. In [36] this was empirically found as the best approach in the smallest applicable case $v = 7$, and the subsequent algebraic analysis in [29] has largely explained this phenomenon.

The smallest rotation-invariant subsets of \mathcal{G}_W admitting a standard rearrangement have size $8(2^n - 1)$ and consist of the rotated copies $r(f + \mathcal{T})$, $r \in \mathbb{F}_{2^n}^\times$, of a single coset $f + \mathcal{T}$ with $f \in \mathcal{R} \setminus \mathcal{T}$. Since there are $2^{n-3} - 1$ such cosets, the total number of choices for \mathcal{A} , being equal to $2^{2^{n-3}-1} - 1$ ($\mathcal{A} = \emptyset$ is omitted), is still growing extremely fast with n .

Relaxing the condition of exact rearrangement, we now ask for maximum-size subsets $\mathcal{N}' \subseteq \mathcal{N}$ of the standard rearrangement of $\{\Gamma_f; f \in \mathcal{A}\}$ into new planes subject to $d_s(\mathcal{N}') \geq 4$. Rotation-invariance of \mathcal{A} reduces this, in general, “global” problem to a “local” problem at one particular point of S , say $P_1 = \mathbb{F}_2(0, 1)$. For the statement of the result note that the group Σ is isomorphic to $\mathbb{F}_{2^n}^\times$ via $((x, y) \mapsto (x, ry)) \mapsto r$.

Lemma 4.1. *Let $\mathcal{A} \subseteq \mathcal{G}_W$ be rotation-invariant and such that $\mathcal{A} \cap \mathcal{R}$ forms a union of nontrivial cosets of \mathcal{T} , \mathcal{N} the corresponding standard rearrangement into new planes, and $\mathcal{N}_r \subseteq \mathcal{N}$ the set of new planes passing through $P_r = \mathbb{F}_2(0, r) \in S$.*

- (i) *If $\mathcal{N}'_1 \subseteq \mathcal{N}_1$ has maximum size M_1 subject to the distance condition $d_s(\mathcal{N}'_1) \geq 4$ then $\mathcal{N}' = \bigcup_{g \in \Sigma} g(\mathcal{N}'_1) \subseteq \mathcal{N}$ has maximum size $M_1(2^n - 1)$ subject to the distance condition $d_s(\mathcal{N}') \geq 4$.*
- (ii) *The totality of subsets $\mathcal{N}' \subseteq \mathcal{N}$ of maximum size $M_1(2^n - 1)$ satisfying $d_s(\mathcal{N}') \geq 4$ is obtained by choosing, independently for each $g \in \Sigma$, subsets $\mathcal{N}_g \subseteq \mathcal{N}_1$ of size M_1 with $d_s(\mathcal{N}_g) \geq 4$ and taking the union $\mathcal{N}' = \bigcup_{g \in \Sigma} g(\mathcal{N}_g)$.*

Therefore, if the number of sets \mathcal{N}'_1 in (i) is t , the total number of choices for \mathcal{N}' in (ii) is equal to t^{2^n-1} .

Proof. Since \mathcal{A} is rotation-invariant, we have $g(\mathcal{N}_1) = \mathcal{N}_r$ for the (unique) element $g \in \Sigma$ that acts as $(x, y) \mapsto (x, ry)$. Hence solving the optimization problem for \mathcal{N}_r is equivalent to solving it for \mathcal{N}_1 . The proof is completed by the observation that $d_s(\mathcal{N}_r, \mathcal{N}_{r'}) \geq 4$ if $r \neq r'$, i.e., new planes meeting S in different points do not conflict.²⁷ \square

²⁷As part of the standard rearrangement, they cannot have a line disjoint from S in common.

The planes in \mathcal{N}_1 satisfy $r\delta(Z) = 1$ in the $U(Z, T, f)$ -representation stated earlier, hence have the form $N = U(Z, P_1, \delta(u, x)/\delta(Z))$ or $N = U(Z, P_1, \delta(u + c, x)/\delta(Z))$.²⁸ The smallest sets \mathcal{A} satisfying the conditions of Lemma 4.1 correspond to a nontrivial coset $u + W$ and hence to a solid of $\text{PG}(\mathbb{F}_{2^n})$ containing W , viz. $T = \langle W, u \rangle$. If u is fixed then, since $\langle Z, u \rangle$ and $\langle Z, u + c \rangle$ account precisely for the 14 planes $\neq W$ in T , we may view this correspondence as a parametrization of the new planes in \mathcal{N}_1 by those planes. If \mathcal{A} is chosen as the largest set satisfying the conditions of Lemma 4.1 (i.e., the set of all rank-3 binomials in \mathcal{G}), all planes of $\text{PG}(\mathbb{F}_{2^n})$ intersecting W in a line are used as parameters.

Definition 4.2. The *collision graph* Γ_W has as its vertices the planes in $\text{PG}(\mathbb{F}_{2^n})$ meeting W in a line. Two vertices $E = \langle Z, u \rangle$ and $E' = \langle Z', u' \rangle$ are adjacent in Γ_W if and only if the new planes $N, N' \in \mathcal{N}_1$ parametrized by E, E' have a point outside S (and hence a line through P_1) in common.

By Lemma 4.1, the graph Γ_W encapsulates all information necessary for the determination of the largest sets of new planes that can be added to the expurgated Gabidulin code without decreasing the subspace distance to 2, for all “starter” sets $\mathcal{A} \cap \mathcal{R} \subset \mathcal{G}_W$ satisfying the assumptions of the lemma: A specific set \mathcal{A} corresponds, via the parametrization $E = \langle Z, u \rangle$, to a certain vertex subgraph of Γ_W , and the maximum-size cocliques of this subgraph yields precisely the largest sets $\mathcal{N}'_1 \subseteq \mathcal{N}_1$ of new planes that can be added in P_1 ; in particular, the number M_1 in Part (i) of the lemma equals the independence number of the subgraph.

If $\mathcal{A} \cap \mathcal{R}$ consists of t cosets of \mathcal{T} , the size of the largest $(v, M, 4; 3)_2$ code \mathcal{C} that can be obtained by this method equals

$$(7) \quad \#\mathcal{C} = 4^n - 8t(2^n - 1) + M_1(2^n - 1) = 4^n + (M_1 - 8t)(2^n - 1).$$

We call the quantity $(M_1 - 8t)(2^n - 1)$ the *net gain* of \mathcal{C} (relative to an LMRD code) and the quantity $M_1 - 8t$ the *local net gain* of \mathcal{C} . The present optimization problem may then be stated as follows.

Rotation-invariant Rearrangement Problem (RRP). Among all $\begin{bmatrix} n \\ 3 \end{bmatrix}_2$, $n = v - 3$, choices for the plane W in $\text{PG}(\mathbb{F}_{2^n})$ and all $2^{2^{n-3}-1} - 1$ choices for a (non-empty) subset $\mathcal{A} \subset \mathcal{G}_W$ satisfying the conditions of Lemma 4.1, determine those which result in the largest (local) net gain for the augmented expurgated Gabidulin code.

The LRMD code bound corresponds to a local net gain of $\begin{bmatrix} n \\ 2 \end{bmatrix}_2 / (2^n - 1) = \frac{2^{n-1}-1}{3}$.

At the first glance, this new rearrangement problem seems just as difficult as the original one, but this is not true. Collisions between new planes through P_1 can be characterized algebraically be a certain invariant of the parametrizing plane E , as observed in [29]. This forces the collision graph Γ_W to have a very special structure, which greatly simplifies the computation of the independence numbers of the relevant subgraphs.

Definition 4.3. The σ -invariant of a plane E in $\text{PG}(\mathbb{F}_{2^n})$ intersecting W in a line Z is defined as

$$\sigma_W(E) = \frac{\delta(E)}{\delta(Z)^3},$$

where $\delta(E)$ denotes the product of all points in E .

²⁸The meaning of c and u is the same as in (6).

The plane invariant $\delta(E)$ is the 3-dimensional analogue of the line invariant $\delta(L)$ and another instance of the Dickson invariants mentioned in Remark 1.

Now we can state and prove the key result on the algebraic characterization of collisions between new planes.

Theorem 4.4. *Two distinct planes E, E' in $\text{PG}(\mathbb{F}_{2^n})$ intersecting W in a line form an edge of the collision graph Γ_W if and only if $\sigma_W(E) = \sigma_W(E')$.*

Proof. Let $E = \langle Z, u \rangle$, $E' = \langle Z', u' \rangle$ with $Z = E \cap W$, $Z' = E' \cap W$, and let N, N' the new planes corresponding to E, E' . Inspecting (6) and using $r\delta(Z) = 1$, we find that the 6 points on N outside S have the form

$$\left(z, \frac{\delta(L)}{\delta(Z)}\right), \quad \left(z, \frac{\delta(L)}{\delta(Z)} + 1\right) \quad \text{with } z \in Z \text{ and } L = \langle z, u \rangle,$$

and similarly for N' . Hence N, N' have a point outside S in common if and only if there exists $z \in Z \cap Z'$ such that, with $L = \langle z, u \rangle$ and $L' = \langle z, u' \rangle$,

$$\left(\frac{\delta(L)}{\delta(Z)}\right)^2 + \frac{\delta(L)}{\delta(Z)} = \left(\frac{\delta(L')}{\delta(Z')}\right)^2 + \frac{\delta(L')}{\delta(Z')}$$

or, equivalently,

$$\frac{\delta(Z)\delta(L)(\delta(Z) + \delta(L))}{\delta(Z)^3} = \frac{\delta(Z')\delta(L')(\delta(Z') + \delta(L'))}{\delta(Z')^3}.$$

But, since $\delta(Z) + \delta(L)$ is the line invariant of the third line in E through z , and similarly for $\delta(Z') + \delta(L')$ (cf. Lemma 3.3 and the remarks preceding it), the latter identity reduces to $z^2\delta(E)/\delta(Z)^3 = z^2\delta(E')/\delta(Z')^3$ and hence to $\sigma(E) = \sigma(E')$.

Conversely, $\sigma(E) = \sigma(E')$ implies that N, N' have a point of the form (z, y) in common for every $z \in Z \cap Z'$. Since Z and Z' intersect (as lines of a projective plane), there exists at least one such z . \square

Remark 3. Using the notation of Theorem 4.4 and its proof, the planes N, N' can form a collision only if $Z \neq Z'$. This fact follows, e.g., from the last part of the proof: $Z = Z'$ implies that N, N' have at least 3 points outside S (one for every $z \in Z$) in common; hence $N = N'$. Alternatively, the fact is a consequence of Theorem 3.4(i), since $Z = Z'$ implies that N, N' correspond to the same $r = \delta(Z)^{-1} = \delta(Z')^{-1}$.

Thus Theorem 4.4 says in particular that the map $E \mapsto \delta(E)$ is one-to-one on the set of planes $E \neq W$ of $\text{PG}(\mathbb{F}_{2^n})$ containing a fixed line $Z \subset W$. This remains true when $E = W$ is included and is a special case of a more general fact; cf. Theorem 5.2 in Section 5.

Theorem 4.4 has the following immediate corollary.

Corollary 1. *The independence number of Γ_W is equal to the number of different values taken by σ_W (i.e., the size of $\text{Im}(\sigma_W)$). Likewise, the independence number of the subgraph of Γ_W corresponding to $\mathcal{A} \subset \mathcal{G}_W$ as in Lemma 4.1 equals the number of different values taken by σ_W on the set of all planes $E \neq W$ that are contained in one of the solids $T \supset W$ corresponding to \mathcal{A} .*

Proof. By Theorem 4.4, Γ_W is a disjoint union of cliques, and the corollary follows. \square

Although not explicitly stated in the corollary, it is clear that all maximum-size cocliques of Γ_W are obtained by selecting for each $y \in \text{Im}(\sigma_W)$ precisely one plane E with $\sigma_W(E) = y$ and that the number of maximum-size cocliques is equal to the product of the multiplicities of all $y \in \text{Im}(\sigma_W)$, and similarly for the subgraphs of Γ_W corresponding to $\mathcal{A} \subset \mathcal{G}_W$.

Another pleasant consequence of Theorem 4.4 is the invariance of the present optimization problem under a fairly large collineation group of $\text{PG}(\mathbb{F}_{2^n})$ acting on the set of all $\begin{bmatrix} n \\ 3 \end{bmatrix}_2$ planes W in $\text{PG}(\mathbb{F}_{2^n})$, which can serve as the first factor of the ambient space V .

Let G be the subgroup of $\text{GL}(\mathbb{F}_{2^n})$ generated by the multiplication maps $x \mapsto rx$, $r \in \mathbb{F}_{2^n}^\times$, and the Frobenius automorphism $\varphi: x \mapsto x^2$. The group G is a Frobenius group with kernel $H = \{x \mapsto rx; r \in \mathbb{F}_{2^n}^\times\} \cong \mathbb{F}_{2^n}^\times$ and complement $K = \langle \varphi \rangle = \text{Aut}(\mathbb{F}_{2^n}/\mathbb{F}_2)$ and has order $\#G = n(2^n - 1)$. It can also be seen as the normalizer of H in $\text{GL}(\mathbb{F}_{2^n})$.

Corollary 2. *If W_1 and W_2 are in the same orbit of G on planes in $\text{PG}(\mathbb{F}_{2^n})$ then the collision graphs Γ_{W_1} and Γ_{W_2} are isomorphic, and the corresponding RRP's are equivalent in the sense that a solution of one problem immediately gives a corresponding solution of the other problem.*

Proof. If $W_2 = rW_1$ for some $r \in \mathbb{F}_{2^n}^\times$ then $E \mapsto rE$ sends the planes meeting W_1 in a line Z to those meeting W_2 in rZ . Clearly we have $\delta(rZ) = r^3\delta(Z)$, $\delta(rE) = r^7\delta(E)$, and hence

$$\sigma_{W_2}(rE) = \frac{\delta(rE)}{\delta(rZ)} = \frac{r^7\delta(E)}{r^3\delta(Z)} = r^4\sigma_{W_1}(E).$$

Together with Theorem 4.4 it follows that $E \mapsto rE$ represents a graph isomorphism $\Gamma_{W_1} \rightarrow \Gamma_{W_2}$. Similarly, if $W_2 = \varphi(W_1) = W_1^2$ then $\sigma_{W_2}(E^2) = \sigma_{W_1}(E)^2$ and $E \mapsto E^2$ represents a graph isomorphism $\Gamma_{W_1} \rightarrow \Gamma_{W_2}$. Since G is generated by the maps $x \mapsto rx$ and φ , the first assertion follows. The second assertion is clear. \square

We close this section with two examples. The first example recalls the construction of a $(7, 301, 4; 3)_2$ code in [36, 29], which was used as the intermediate step in an alternative construction of a currently best known $(7, 329, 4; 3)_2$ code in [36, 29], following the original discovery of such a code in [11].

Example 2 ($v = 7$). Here $n = 4$ and we represent \mathbb{F}_{16} as $\mathbb{F}_2(\xi)$ with $\xi^4 + \xi + 1 = 0$. Since all planes in $\text{PG}(\mathbb{F}_{16})$ are conjugate under multiplication (by Singer's Theorem), Corollary 2 implies that we can choose W freely; for convenience, we take $W = \{x \in \mathbb{F}_{16}; \text{Tr}(x) = 0\} = \{1, \xi, \xi^2, \xi^4, \xi^5, \xi^8, \xi^{10}\}$. Since $\mathcal{A} = \mathcal{R} \setminus \mathcal{T} = \{ux^2 + u^2x; \text{Tr}(u) = 1\}$ is a single coset in this case, there is only one choice for the rotation-invariant set \mathcal{A} in the RRP; \mathcal{A} consists of the 120 rank-3 binomials $r(ux^2 + u^2x)$ with $u \in \mathbb{F}_{16} \setminus W$, $r \in \mathbb{F}_{16}^\times$.

The next step is to compute the σ -invariants of the 14 planes $E \neq W$ in $\text{PG}(\mathbb{F}_{16})$. If $E = aW$, $a \in \mathbb{F}_{16} \setminus \mathbb{F}_2$, then $\delta(E) = a^7\delta(W) = a^7$ and $Z = E \cap W = W \cap aW = \{x \in \mathbb{F}_{16}; x^8 + x^4 + x^2 + x = a^{-8}x^8 + a^{-4}x^4 + a^{-2}x^2 + a^{-1}x = 0\}$. Eliminating, we find $(1 + a^4)x^4 + (1 + x^6)x^2 + (1 + a^7)x = 0$, $\delta(Z) = \frac{1+a^7}{1+a^4}$,

$$\sigma(aW) = \frac{a^7(1+a^4)^3}{(1+a^7)^3} = \frac{a^7(1+a^4)(1+a^8)}{(1+a^7)(1+a^{14})} = \frac{1+a^4}{1+a^{14}} = a + a^2 + a^3 + a^4,$$

a special case of [29, Lemma 7]. This tells us that $\sigma(aW) = 1$ for $a \in \{\xi^3, \xi^6, \xi^9, \xi^{12}\}$ (the 5-th primitive roots of unity in \mathbb{F}_{16}^\times) and, not difficult to verify from the representation $\sigma(aW) = a(1+a)^3$, that $E \mapsto \sigma(E)$ maps the remaining 10 planes $E \neq W$ bijectively onto $\mathbb{F}_{16}^\times \setminus \{1, \xi^3, \xi^6, \xi^9, \xi^{12}\}$.

It follows that the collision graph Γ consists of a complete graph K_4 (formed by $\xi^3W, \xi^6W, \xi^9W, \xi^{12}W$) and 10 isolated vertices. The independence number of Γ is 11, and we can add locally at P_1 the 10 new planes parametrized by aW , a not a 3rd power in \mathbb{F}_{16}^\times , to the expurgated Gabidulin code, and exactly one of the four new planes parametrized by $\xi^{3i}W$ ($1 \leq i \leq 4$). Finally, rotating through the 15 points of S , we obtain 4^{15} different extensions of the expurgated Gabidulin code to a plane subspace code of size $256 - 120 + 15 \times 11 = 301$. Exactly 4 of these are rotation-invariant. Explicit representations of the new codewords through P_1 may be obtained by writing $aW = \langle Z, u \rangle$ with $Z = W \cap aW$ and evaluating $N = U(Z, P_1, \delta(u, x)/\delta(Z))$ explicitly.²⁹

Our second example provides a solution of the RRP in the smallest open case.

Example 3 ($v = 8$). Here we have $n = 5$ and the relevant extension field of \mathbb{F}_2 is $\mathbb{F}_{32} = \mathbb{F}_2[\alpha]$ with $\alpha^5 + \alpha^2 + 1 = 0$. In this case the group G , of order $5 \cdot 31 = 155 = [5]_2$, acts (sharply) transitive of the set of all planes in $\text{PG}(\mathbb{F}_{32}) \cong \text{PG}(4, \mathbb{F}_2)$. Hence, by Corollary 2 it suffices to consider one particular plane W , which we can take as $W = \langle 1, \alpha, \alpha^2 \rangle$, determining the ambient space $V = W \times \mathbb{F}_{32}$.

The number of nontrivial cosets in $\mathcal{R}/\mathcal{T} \cong \mathbb{F}_{32}/W$ is 3, so that $2^3 - 1 = 7$ different coset combinations need to be considered for \mathcal{A} . The 14 new planes in \mathcal{N}_1 corresponding to a minimal choice of \mathcal{A} (1 coset) are represented by the 14 planes $\neq W$ in one of the solids $T_1, T_2, T_3 \supset W$ in $\text{PG}(\mathbb{F}_{32})$. Using the computer algebra system SageMath (www.sagemath.org), we have found that σ_W takes 11 distinct values on each of the three 14-sets of planes E contained in a fixed T_i and that the values of multiplicity > 1 in each case are the same, viz. $\alpha^{23}, \alpha^{25}, \alpha^{28}$, all of multiplicity 2.³⁰ This implies $\#\text{Im}(\sigma_W) = 27$ (the 4 “missing” values are $\alpha^4, \alpha^5, \alpha^{21}, \alpha^{30}$) and that the independence number of any subgraph of Γ_W involving 1, 2, 3 cosets is 11, 19, and 27 respectively.³¹ Since $11 - 8 = 19 - 2 \cdot 8 = 27 - 3 \cdot 8 = 3$, the local net gain when using $t \in \{1, 2, 3\}$ cosets is always 3, a constant independent of t , and the global net gain is $3 \cdot 31 = 93$. Thus the largest subspace codes obtained by rotation-invariant rearrangement from the Gabidulin code have size $1024 + 93 = 1117$. Since the largest known $(8, M, 4; 3)_2$ code has size 1326 (at the time of writing this article, cf. [10]), these codes are not particularly good. However, they can be further extended by planes meeting $S = \{0\} \times \mathbb{F}_{32}$ in a line; cf. Section 7.

Although the machinery developed so far is sufficient for a complete solution of the RRP for $v = 8$ (with the aid of a computer), the computational complexity

²⁹Since $N = \langle \Gamma_f, P_1 \rangle$ for $f: Z \rightarrow \mathbb{F}_{16}, x \mapsto \delta(u, x)/\delta(Z)$, it suffices to determine the graphs of these linear maps.

³⁰Since $\alpha^{23} + \alpha^{25} + \alpha^{28} = 0$, these “collision values” form a line in $\text{PG}(\mathbb{F}_{32})$.

³¹Thus the whole graph Γ_W , corresponding to all 3 cosets, has independence number 27 and consists of 24 isolated vertices and 3 cliques of size 6, which intersect the three 14-sets in a 2-set.

of the presently used naive method for determining the best coset combination (“exhaustive search”) is prohibitive for only slightly larger values of v .³²

5. DICKSON INVARIANTS, SUBSPACE POLYNOMIALS AND ALL THAT

In this section we develop the machinery that is needed to understand the subsequent analysis of the collision graphs Γ_W and their condensed variants, called *collision matrices*, which provide all essential information about the clique sizes in Γ_W . The relevant background can be found in the seminal work of ORE on linearized polynomials [39] and to some extent in BERLEKAMP’s book [5, Ch. 11]. As usual we will restrict ourselves to the ground field \mathbb{F}_2 , although everything can be generalized with only little more effort to \mathbb{F}_q .

A convenient starting point is the following 2-analogue of the well-known Vandermonde determinant evaluation due to E. H. MOORE [38], which holds as an identity in the polynomial ring $\mathbb{F}_2[X_1, \dots, X_k]$:

$$(8) \quad \delta(X_1, \dots, X_k) = \begin{vmatrix} X_1 & X_2 & \dots & X_k \\ X_1^2 & X_2^2 & \dots & X_k^2 \\ X_1^{2^2} & X_2^{2^2} & \dots & X_k^{2^2} \\ \vdots & \vdots & & \vdots \\ X_1^{2^{k-1}} & X_2^{2^{k-1}} & \dots & X_k^{2^{k-1}} \end{vmatrix} = \prod_{\lambda \in \mathbb{F}_2^k \setminus \{0\}} (\lambda_1 X_1 + \dots + \lambda_k X_k).$$

This identity can be proved using induction on k and

$$(9) \quad \delta(X_1, \dots, X_k) = \delta(X_1, \dots, X_{k-1}) \prod_{\lambda \in \mathbb{F}_2^{k-1}} (X_k + \lambda_1 X_1 + \dots + \lambda_{k-1} X_{k-1}).$$

The latter identity is obtained in the same way as for the ordinary Vandermonde determinant by viewing the determinant in (8) as a polynomial in X_k over the rational function field $\mathbb{F}_2(X_1, \dots, X_{k-1})$ and determining its zeros.

Now let U be a k -dimensional \mathbb{F}_2 -subspace of \mathbb{F}_{2^n} with basis β_1, \dots, β_k . Replacing k by $k+1$ in (9) and making appropriate substitutions, we obtain the identity

$$(10) \quad \begin{aligned} \prod_{u \in U} (X + u) &= \prod_{\lambda \in \mathbb{F}_2^k} (X + \lambda_1 \beta_1 + \dots + \lambda_k \beta_k) \\ &= \frac{\delta(\beta_1, \dots, \beta_k, X)}{\delta(\beta_1, \dots, \beta_k)} = \sum_{i=0}^k a_i X^{2^i} \in \mathbb{F}_{2^n}[X]. \end{aligned}$$

The last step uses Laplace expansion of the determinant in (8) along the last column and shows that a_i is equal to the quotient of a certain $k \times k$ determinant involving the powers $\beta_j^{2^t}$, $t \in \{0, \dots, k\} \setminus \{i\}$, and $\delta(\beta_1, \dots, \beta_k)$.

The polynomial $s_U(X) = \prod_{u \in U} (X + u)$, which is monic and has the elements of U as roots of multiplicity 1, is known as the *subspace polynomial* associated with U . From the previous computation we have that $s_U(X)$ is a monic linearized polynomial (2-polynomial) of symbolic degree k . Conversely, a monic 2-polynomial

³²For example, there is absolutely no way to settle the case $v = 11$ in this manner in a reasonable time, since the number of coset combinations that must be explored is $2^{2^{v-6}-1} - 1 = 2^{31} - 1 = 2147483647$.

in $\mathbb{F}_{2^n}[X]$ is a subspace polynomial of some \mathbb{F}_2 -subspace of \mathbb{F}_{2^n} if it splits into linear factors and the coefficient a_0 of X is $\neq 0$.³³

The coefficients of $s_U(X)$ will be called *Dickson invariants* of U and denoted by $\delta_i(U) = a_{k-i}$.³⁴ For the last Dickson invariant $\delta_k(U) = \prod_{u \in U \setminus \{0\}} u = \delta(\beta_1, \dots, \beta_k)$, the coefficient of X in $s_U(X)$, we usually write simply $\delta(U)$.³⁵

The set L_n of 2-polynomials in $\mathbb{F}_{2^n}[X]$ is closed with respect to addition and composition of polynomials (also called “symbolic multiplication”), defined by $f(X) \circ g(X) = f(g(X))$, and forms a ring $(L_n, +, \circ)$. The ring L_n is non-commutative (except for $n = 1$) and isomorphic to the skew polynomial ring $\mathbb{F}_{2^n}[Y; \varphi]$ via $\sum a_i X^{2^i} \mapsto \sum a_i Y^i$. It is thus quite easy to work with.³⁶ One can show that L_n has no zero divisors and admits one-sided analogues of the Euclidean Algorithm for symbolic division. The center of L_n consists of all polynomials of the form $c_0 X + c_1 X^{2^n} + c_2 X^{4^n} + \dots$ (“ 2^n -polynomials”). In particular, $X^{2^n} + X$ is central and $(X^{2^n} + X) = L_n(X^{2^n} + X) = (X^{2^n} + X)L_n$ is a two-sided ideal in L_n .³⁷

The computation of subspace polynomials is facilitated by the following symbolic factorization into linear factors:

$$(11) \quad s_U(X) = (X^2 + s_{k-1}(\beta_k)X) \circ \dots \circ (X^2 + s_1(\beta_2)X) \circ (X^2 + \beta_1 X),$$

where $s_i(X) = s_{\langle \beta_1, \dots, \beta_i \rangle}(X)$. This identity follows by induction on k from $s_U(X) = (X^2 + s_{U'}(\beta)X) \circ s_{U'}(X)$, valid for any incident pair $U' \subset U$ of subspaces of \mathbb{F}_{2^n} with $\dim(U') = k-1$, $\dim(U) = k$, and for any $\beta \in U \setminus U'$. The latter can be proved as follows: Since $U = U' \uplus (\beta + U')$, we have

$$(12) \quad \begin{aligned} s_U(X) &= s_{U'}(X) s_{U'}(X + \beta) \\ &= s_{U'}(X) (s_{U'}(X) + s_{U'}(\beta)) \\ &= s_{U'}(X)^2 + s_{U'}(\beta) s_{U'}(X) \\ &= (X^2 + s_{U'}(\beta)X) \circ s_{U'}(X), \end{aligned}$$

as desired. In the special case $U = \mathbb{F}_{2^n}$ we obtain a symbolic linear factorization of $X^{2^n} + X$, which can be seen as a noncommutative analogue of the ordinary factorization of $X^n + 1$. Since such factorizations are in 1-1 correspondence with ordered bases of U , they are highly non-unique.³⁸

³³Thus every monic 2-polynomial with $a_0 \neq 0$ becomes a subspace polynomial when considered over its splitting field.

³⁴The usual Dickson invariants studied in Modular Invariant Theory (and here specialized to the case $q = 2$) are the polynomial counterparts $\delta_i^{(k)}(X_1, \dots, X_k) \in \mathbb{F}_2[X_1, \dots, X_k]$, which can be obtained in the same way as the coefficients of the “generic” subspace polynomial $\prod_{\lambda \in \mathbb{F}_2^k} (X + \lambda_1 X_1 + \dots + \lambda_k X_k) \in \mathbb{F}_2(X_1, \dots, X_k)[X]$. see [12, 16, 43, 47]. The indexing of $\delta_i, \delta_i^{(k)}$ follows the convention used for the elementary symmetric polynomials; note, however, that the degree of $\delta_i^{(k)}$ is not i but $2^k - 2^{k-i}$.

³⁵This is compatible with the notation used in the cases $k = 2, 3$, which have already been considered.

³⁶The ring $\mathbb{F}_{2^n}[Y; \varphi]$ differs from the ordinary polynomial ring $\mathbb{F}_{2^n}[Y]$ by the law $Ya = \varphi(a)Y = a^2 Y$, which leads to a formula for the coefficients of $f(X) \circ g(X)$ similar to ordinary polynomial multiplication except that the coefficients of $g(X)$ are “twisted” by powers of the Frobenius automorphism. In the special case $n = 1$ (which is not of interest to us here) this ring, and hence L_1 as well, is commutative and isomorphic to $\mathbb{F}_2[Y]$.

³⁷This fact is needed below.

³⁸For example, the number of different symbolic linear factorizations of $X^{2^n} + X$ in L_n is equal to $(2^n - 1)(2^n - 2) \dots (2^n - 2^{n-1}) = \# \text{GL}(n, \mathbb{F}_2)$.

An important aspect of the theory is the interplay between 2-polynomials and \mathbb{F}_2 -linear endomorphisms of extension fields \mathbb{F}_{2^n} . Every such endomorphism is represented by a unique 2-polynomial in L_n of symbolic degree $< n$, and composition of endomorphisms corresponds to symbolic multiplication of 2-polynomials. Using these facts it is not hard to see that $\text{End}(\mathbb{F}_{2^n}/\mathbb{F}_2) \cong L_n/(X^{2^n} + X)$.³⁹ A subspace polynomial represents a particular \mathbb{F}_2 -linear map $\mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^n}$, $x \mapsto s_U(x)$ with kernel U . Its image will be called the *opposite subspace* of U and denoted by U° .

The opposite subspace U° is characterized by the identity $s_{U^\circ}(X) \circ s_U(X) = X^{2^n} + X$, which follows from the observation that both sides represent the zero map in $\text{End}(\mathbb{F}_{2^n}/\mathbb{F}_2)$ (and considering symbolic degrees).⁴⁰ Since $X^{2^n} + X$ is in the center of the ring L_n , we also have $s_U(X) \circ s_{U^\circ}(X) = X^{2^n} + X$,⁴¹ and hence $U^{\circ\circ} = U$. Further, we see that the polynomials in $\mathbb{F}_{2^n}[X]$ that represent subspace polynomials of \mathbb{F}_2 -subspaces of \mathbb{F}_{2^n} are precisely the monic symbolic divisors of $X^{2^n} + X$ in L_n (on either side),⁴² and symbolic factors of subspace polynomials (on either side) are again subspace polynomials. But symbolic products of subspace polynomials are not necessarily subspace polynomials, in view of the extra conditions imposed on the factors in (11).⁴³

Subspace polynomials are thus analogous to generator polynomials of cyclic codes, and the subspace polynomial of U° is the 2-analogue, or non-commutative analogue, of the check polynomial of a cyclic code. But the analogy goes still further, as we will see in a moment.

Apart from the opposite subspace U° , the following subspaces associated with U will be needed later: The *orthogonal subspace* of U is $U^\perp = \{y \in \mathbb{F}_{2^n}; \text{Tr}(xy) = 0 \text{ for all } x \in U\}$, and the *adjoint subspace* of U is the subspace U^* generated by

$$(13) \quad \frac{\delta(\beta_2, \dots, \beta_k)}{\delta(\beta_1, \dots, \beta_k)}, \frac{\delta(\beta_1, \beta_3, \dots, \beta_k)}{\delta(\beta_1, \dots, \beta_k)}, \dots, \frac{\delta(\beta_1, \dots, \beta_{k-1})}{\delta(\beta_1, \dots, \beta_k)}.$$

The definition of U^* does not depend on the chosen basis β_1, \dots, β_k of U , as is easily shown by multilinear expansion. Moreover, one can show that the elements in (13) are linearly independent (cf. the proof of Theorem 5.2 below) and hence $\dim(U^*) = \dim(U)$.

ORE [39] has defined U^* in a different way as the space whose square $(U^*)^2 = \varphi(U^*)$ is the set of roots of the 2-polynomial $s_U(X)^* = \sum_{i=0}^k (a_i X)^{2^{k-i}}$, a 2-analogue of the reciprocal polynomial $X^{\deg p} p(X^{-1})$ associated with an ordinary polynomial $p(X)$, and derived the basis of U listed in (13). ORE [39] has also shown that the nonzero elements in U^* are precisely the elements $A^{-1} \in \mathbb{F}_{2^n}$ for which $X^2 + AX$

³⁹Using again skew polynomial rings, this can be extended to $\text{End}(\mathbb{F}_{2^n}/\mathbb{F}_2) \cong L_n/(X^{2^n} + X) \cong \mathbb{F}_{2^n}[Y; \varphi]/(Y^n + 1)$.

⁴⁰Berlekamp [5, Th. 11.35] denotes $s_{U^\circ}(X)$ by $s_U(X)^*$, which conflicts with our (ORE's) notation for the adjoint subspace or polynomial; see below.

⁴¹This is an instance of the following general fact about noncommutative integral domains R : $ab \in Z(R)$ implies $ab = ba$, which is immediate from $a(ab) = (ab)a = a(ba)$.

⁴²If $X^{2^n} + X = a(X) \circ b(X)$ then $x \mapsto a(b(x))$ is the zero map in $\text{End}(\mathbb{F}_{2^n}/\mathbb{F}_2)$, and hence the dimensions of the kernels of $x \mapsto a(x)$ and $x \mapsto b(x)$ must be equal to their symbolic degrees, i.e., $a(X)$ and $b(X)$ must be subspace polynomials (provided they are monic).

⁴³More precisely, $s_U(X) \circ s_V(X)$ is a subspace polynomial iff $U \subseteq V^\circ$.

is a symbolic left factor of $s_U(X)$.⁴⁴ This follows from

$$(14) \quad s_U(X) = (X^2 + AX) \circ s_V(X) = s_V(X)^2 + As_V(X),$$

which implies $A^{-1} = \frac{\delta(V)}{\delta(U)}$, together with the observation that the nonzero elements of U^* have the form $\frac{\delta(V)}{\delta(U)}$ for some subspace $V \subset U$ of codimension 1; cf. Theorem 5.2 below.

The following lemma relates the three subspaces associated with U and will be needed in Section 6.

Lemma 5.1. *For any subspace of \mathbb{F}_{2^n} we have $(U^*)^2 = (U^\circ)^\perp$.*

Proof. Dividing (14) by A^2 gives $s_U(X)/A^2 = (s_V(X)/A)^2 + s_V(X)/A$. Hence, by Hilbert's Satz 90, $\text{Tr}(s_U(x)/A^2) = 0$ for all $x \in \mathbb{F}_{2^n}$. Since U° is the image of $x \mapsto s_U(x)$, this says $(U^*)^2 \subseteq (U^\circ)^\perp$. Since $\dim(U^*) = \dim(U) = \dim((U^\circ)^\perp)$, the result follows. \square

Since $(U^*)^2$ is the subspace associated with the “reciprocal” subspace polynomial $a_0^{-2^k} s_U(X)^*$, Lemma 5.1 provides a nice 2-analogue of the well-known fact that the dual code C^\perp of a cyclic code C is generated by the reciprocal of the check polynomial of C .

The following properties of the map $U \mapsto \delta(U)$ will play a crucial role in Section 6, and we state them as a theorem.

Theorem 5.2. *Let U be a k -subspace of \mathbb{F}_{2^n} ,*

- (i) *$V \mapsto \delta(V)$ maps the $(k+1)$ -subspaces of \mathbb{F}_{2^n} containing U bijectively onto the 1-subspaces of the space $\delta(U)U^\circ$. The induced map from $\text{PG}(\mathbb{F}_{2^n})/U$ to $\text{PG}(\delta(U)U^\circ)$ is a collineation.*
- (ii) *$V \mapsto \delta(V)$ maps the $(k-1)$ -subspaces of \mathbb{F}_{2^n} contained in U bijectively onto the 1-subspaces of $\delta(U)U^*$. The induced map from $\text{PG}(U)$ to $\text{PG}(\delta(U)U^*)$ is a correlation.*

Note that our earlier Lemma 3.3 is precisely the case $k = 3$ of Part (ii).

Proof. (i) From either (10) or (12) we have $\delta(V) = s_U(\beta)\delta(U)$ for any $(k+1)$ -subspace $V \supset U$ and any $\beta \in V \setminus U$. As β varies over $V \setminus U$, $s_U(\beta)$ varies over $U^\circ \setminus \{0\}$. This proves the first assertion. The second assertion follows from linearity of $\beta \mapsto s_U(\beta)$.

(ii) Let V be a $(k-1)$ -subspace of U with basis v_1, \dots, v_{k-1} . Using multilinear expansion, $\delta(V)/\delta(U) = \delta(v_1, \dots, v_{k-1})/\delta(U)$ can be expressed as an \mathbb{F}_2 -linear combination of the elements in (13), showing that $\delta(V) \in \delta(U)U^*$. The coefficients μ_i of this linear combination are easily seen to be the minors of order $k-1$ of the matrix $(\lambda_{ij}) \in \mathbb{F}_2^{k \times (k-1)}$ determined by $v_j = \sum_{i=1}^k \lambda_{ij} \beta_i$. Now it is well-known that for any $(\mu_1, \dots, \mu_k) \in \mathbb{F}_2^k \setminus \{0\}$ there exists a corresponding matrix (λ_{ij}) of rank $k-1$ having μ_i as their order $k-1$ minors.⁴⁵ Hence any element of U^* has the form $\delta(V)/\delta(U)$ and $V \mapsto \delta(V)$ maps onto $\delta(U)U^*$. The elements in (13) are linearly

⁴⁴Compare this with the obvious fact that the $X^2 + aX$ is a symbolic right factor of U iff $a \in U$. More generally, we have $V \subseteq U$ ($V^* \subseteq U^*$) iff $s_V(X)$ is a symbolic right (respectively, left) factor of $s_U(X)$.

⁴⁵Essentially this amounts to the fact that any nonzero vector in \mathbb{F}_2^k can be completed to an invertible $k \times k$ matrix.

independent. (A dependency relation would yield a subspace V with $\delta(V) = 0$ by what we have just shown; this is impossible.) Hence $\dim(U^*) = \dim(U)$ and $V \mapsto \delta(V)$ maps the $(k-1)$ -subspaces of U bijectively onto $\delta(U)U^*$. Finally, the correlation property follows from Part (i): If $V_0 \subset U$ has dimension $k-2$ then the $(k-1)$ -subspaces between V_0 and U are mapped to a 2-subspace of $\delta(V_0)V_0^\circ$, which must be contained in $\delta(U)U^*$. \square

Theorem 5.2 has the following rather curious corollary.

Corollary 3. *The k -subspaces $U \subseteq \mathbb{F}_{2^v}$ with fixed last Dickson invariant $\delta(U) = a$, $a \in \mathbb{F}_{2^v}^\times$, form a subspace code $\mathcal{C}(a)$ with minimum distance at least 4.*

Proof. If $U, V \in \mathcal{C}$ satisfy $d_s(U, V) = 2$ then $\dim(U \cap V) = k-1$, $\dim(U+V) = k+1$. Now either Part (i) of the theorem applied to $W = U \cap V$, or Part (ii) applied to $W = U + V$ yields a contradiction. \square

By Corollary 3, the set of k -subspaces of \mathbb{F}_2^v is partitioned into $2^v - 1$ (possibly empty) subspace codes of minimum distance ≥ 4 . Viewed as single codes, these are not very interesting, since they are too small. In the case $k = 3$ the largest of these codes has guaranteed size

$$\#\mathcal{C}(a) \geq \frac{1}{2^v - 1} \begin{bmatrix} v \\ 3 \end{bmatrix}_2 = \frac{(2^{v-1} - 1)(2^{v-2} - 1)}{21} \approx \frac{8}{21} 2^{2(v-3)},$$

which is considerably smaller than the size of the corresponding Gabidulin codes. A computational study of combinations of several codes $\mathcal{C}(a)$ has not produced anything of value. Nevertheless, the corollary and its ramifications deserve further research. Links between subspace polynomials and subspace codes are also studied in [2], with emphasis on the case of cyclic subspace codes. The codes of Corollary 3 tend to be transversal to the corresponding Singer orbits and are decidedly non-cyclic. On the other hand, the gap theorem in [2, Cor. 2] and our corollary are similar—in both cases certain Dickson invariants are prescribed.

We close this section with some simple examples of subspace polynomial computations. About point polynomials $s_P(X) = X^2 + aX$, $P = \mathbb{F}_2 a$, there is not much to say. Line polynomials $s_L(X)$, $L = \langle a, b \rangle$ are easily computed using (11) and have the form $s_L(X) = (X^2 + (b^2 + ab)X) \circ (X^2 + aX) = (X^2 + aX)^2 + (b^2 + ab)(X^2 + aX) = X^4 + (a^2 + ab + b^2)X^2 + (ab^2 + a^2b)X$, recovering $\delta_2(L) = \delta(L) = ab^2 + a^2b = ab(a+b)$ and showing that $\delta_1(L) = a^2 + ab + b^2$. We have $\delta_1(L) = 0$ iff $(a/b)^3 = 1$, a property characterizing the lines of the standard line spread in $\text{PG}(\mathbb{F}_{2^n})$, n even (since these lines have the form $a\mathbb{F}_4^\times$). The most prominent example is the well-known $s_{\mathbb{F}_4}(X) = X^4 + X$. Sometimes it is useful to express $\delta_1(L)$, $\delta_2(L)$ in terms of each other, for which we note that $\delta_1(L)\delta_2(L) = ab(a^3 + b^3) = ab^4 + a^4b$. Often subspace polynomial computations can be simplified by taking the action of the multiplicative group, $U \mapsto rU$ for $r \in \mathbb{F}_{2^n}^\times$ and of the Frobenius automorphism ($U \mapsto U^2 = \{u^2, u \in U\}$) into account. This is illustrated in the final example of this section.

Example 4. We compute the subspace polynomials for all lines and planes in \mathbb{F}_{16} . One example of a plane polynomial is the well-known polynomial $s_W(X) = X^8 + X^4 + X^2 + X$ of the trace-zero plane $W = \{x \in \mathbb{F}_{16}; \text{Tr}(x) = 0\}$. Any further plane has the form rW for a unique $r \in \mathbb{F}_{16}^\times$, and the corresponding polynomial is

$$s_{rW}(X) = r^8 s_W(r^{-1}X) = X^8 + r^4 X^4 + r^6 X^2 + r^7 X.$$

For line polynomials $s_L(X) = X^4 + a_1X^2 + a_0X$ we use the formulas

$$\begin{aligned} s_{rL}(X) &= r^4 s_L(r^{-1}X) = X^4 + r^2 a_1 X^2 + r^3 a_0 X, \\ s_{L^2}(X) &= \prod_{u \in L} (X + u^2) = X^4 + a_1^2 X^2 + a_0^2 X, \end{aligned}$$

i.e. $L \mapsto rL$ and $L \mapsto L^2$ correspond to $(a_0, a_1) \mapsto (r^3 a_0, r^2 a_1)$ and $(a_0, a_1) \mapsto (a_0^2, a_1^2)$. Using $\mathbb{F}_{16} = \mathbb{F}_2(\xi)$ with $\xi^4 + \xi + 1 = 0$ and $\omega = \xi^5$, we have $\mathbb{F}_{16}^\times = \{\xi^i; 0 \leq i \leq 14\}$ and $\mathbb{F}_2(\omega) = \mathbb{F}_4$ inside \mathbb{F}_{16} . The lines of the standard spread of $\text{PG}(\mathbb{F}_{16})$ have polynomials $s_{\xi^i \mathbb{F}_4}(X) = X^4 + \xi^{3i} X$ ($0 \leq i \leq 4$). The remaining 30 lines in $\text{PG}(\mathbb{F}_{16})$ can be obtained from a single line L as rL or rL^2 . Since with rL the coefficient a_1 in $s_{rL}(X)$ “rotates” through all of \mathbb{F}_{16}^\times , there exist exactly two line polynomials of the form $X^4 + X^2 + a_0$, which must be $X^4 + X^2 + \omega X$ and $X^4 + X^2 + \omega^2 X$ (since their conjugates under $x \mapsto x^2$ are line polynomials as well). Hence the remaining 30 line polynomials are $X^4 + r^2 X^2 + r^3 \omega^i X$ with $i \in \{1, 2\}$ and $r \in \mathbb{F}_{16}^\times$, and it only remains to identify the line L behind $X^4 + X^2 + \omega X = X(X^3 + X + \omega)$ by factoring this polynomial. It turns out that $L = \{\xi^{10}, \xi^{11}, \xi^{14}\} = \xi^{10}\langle 1, \xi \rangle$.⁴⁶

6. CONTINUATION OF THE ANALYSIS

Our first goal in this section is to determine the set of multiple values of σ_W . The planes E in $\text{PG}(\mathbb{F}_{2^n})$ meeting W in a line fall into 7 classes according to their intersection $Z = E \cap W$. By Theorem 5.2(i), the restriction of $\sigma_W: E \mapsto \delta(E)/\delta(Z)^3$ to such a class is one-to-one with image $\delta(Z)^{-2}Z^\circ$, provided we include $\delta(W)/\delta(Z)^3$ in the image. We will refer to the 7 values $\delta(W)/\delta(Z)^3$, $Z \subset W$ a line, as the “missing values” (or “missing points”) of σ_W .⁴⁷

For $x \in \mathbb{F}_{2^n} \setminus Z$ we have $\delta(\langle Z, x \rangle) = s_Z(x)\delta(Z)$ and hence

$$\sigma_W(\langle Z, x \rangle) = \frac{s_Z(x)}{\delta(Z)^2}.$$

For $x, y \in \mathbb{F}_{2^n}$ we set

$$(15) \quad \langle x, y \rangle_Z = \text{Tr} \left(\frac{s_Z(x)y^2}{\delta(Z)^2} \right).$$

Lemma 6.1. *For any line Z in $\text{PG}(\mathbb{F}_{2^n})$ the map $\mathbb{F}_{2^n} \times \mathbb{F}_{2^n} \rightarrow \mathbb{F}_2$, $(x, y) \mapsto \langle x, y \rangle_Z$ is a symmetric \mathbb{F}_2 -bilinear form with radical Z and associated quadratic form $x \mapsto \langle x, x \rangle_Z = \text{Tr}(a_1 x^4 / a_0^2)$, where $a_0 = \delta(Z)$ and $a_1 = \delta_1(Z)$.*

Proof. Bilinearity is clear. Further, we have

$$(16) \quad \frac{s_Z(x)y^2}{\delta(Z)^2} = \frac{(x^4 + a_1 x^2 + a_0 x)y^2}{a_0^2} = \frac{x^4 y^2}{a_0^2} + \frac{a_1 x^2 y^2}{a_0^2} + \frac{xy^2}{a_0}.$$

Now note that the trace of the sum in (16) does not change if we conjugate the three summands individually (!) with powers of the Frobenius automorphism.⁴⁸

⁴⁶Perhaps the easiest way to find L is to compute the line polynomial of $\langle 1, \xi \rangle = \{1, \xi, \xi^4\}$, which is $X^4 + \xi^{10} X^2 + \xi^5 X$, and rotate by $r = (\xi^{10})^{-1/2} = \xi^{10}$.

⁴⁷The missing values are counted with their multiplicities. They are not necessarily all different, and they can still be in the image of σ_W ; cf. Theorem 6.2 for the details.

⁴⁸We have found this “trace trick” useful on several occasions.

Expressing everything in terms of x^2 , we get

$$\begin{aligned}\langle x, y \rangle_Z &= \text{Tr} \left(\frac{x^2 y}{a_0} + \frac{a_1 x^2 y^2}{a_0^2} + \frac{x^2 y^4}{a_0^2} \right) \\ &= \text{Tr} \left(\frac{x^2 (a_0 y + a_1 y^2 + y^4)}{a_0^2} \right) = \langle y, x \rangle_Z.\end{aligned}$$

Since the trace is nondegenerate and $s_Z(x) = 0$ iff $x \in Z$, the radical of $\langle x, y \rangle_Z$ must be Z . Finally, substituting $y = x$ into (16) turns the third summand into a conjugate of the first, giving $\langle x, x \rangle_Z = \text{Tr}(a_1 x^4 / a_0^2)$. \square

Theorem 6.2. (i) The spaces $\delta(Z)^{-2}Z^\circ$, $Z \subset W$ a line, mutually intersect in $(W^2)^\perp$, and hence account for all $(n-2)$ -subspaces of \mathbb{F}_{2^n} containing $(W^2)^\perp$.
(ii) The set of multiple values of σ_W is precisely the subspace $(W^2)^\perp$.⁴⁹

Before proving the theorem we note the following consequence of Part (i): $y \in \mathbb{F}_{2^n}^\times$ is not in the image of σ_W iff y is a missing point of σ_W and $y \notin (W^2)^\perp$. This shows $\#\text{Im}(\sigma_W) = 2^n - 1 - (7 - \mu)$, where μ denotes the number of missing points contained in $(W^2)^\perp$. Since $\mu \leq 7$, σ_W is “almost” surjective for large n .

Proof. (i) For $y \in \mathbb{F}_{2^n}^\times$ consider the seven equations

$$(17) \quad \frac{s_Z(x)}{\delta(Z)^2} = y, \quad Z \text{ a line in } W.$$

For any particular Z , (17) is solvable iff $y \in \delta(Z)^{-2}Z^\circ$. On the other hand, we will show that (17) is solvable iff $y \in (Z^2)^\perp$, thereby establishing $\delta(Z)^{-2}Z^\circ = (Z^2)^\perp$. Since $\bigcap \{(Z^2)^\perp; Z \subset W \text{ a line}\} = (W^2)^\perp$ and $\dim(W^2)^\perp = n - 3$, (i) then follows.

Using Lemma 6.1, (17) implies $\text{Tr}(yz^2) = \langle x, z \rangle_Z = \langle z, x \rangle_Z = 0$ for all $z \in Z$ and thus $y \in (Z^2)^\perp$. Conversely, suppose $y \in (Z^2)^\perp$ and $c \in \mathbb{F}_{2^n}$ is such that $\text{Tr} \left(\frac{s_Z(x)c^2}{\delta(Z)^2} \right) = \langle x, c \rangle_Z = 0$ for all $x \in \mathbb{F}_{2^n}$. Then $c \in Z$, the radical of $\langle \cdot, \cdot \rangle_Z$, and hence $\text{Tr}(yc^2) = 0$. The non-degeneracy of the trace bilinear form now implies that (17) has a solution, completing the proof of (i).

Let us remark that the identity $\delta(Z)^{-2}Z^\circ = (Z^2)^\perp$, which also shows that $Z \mapsto \delta(Z)^{-2}Z^\circ$ defines a correlation from $\text{PG}(W)$ to $\text{PG}(\mathbb{F}_{2^n}/(W^2)^\perp)$, can alternatively be derived from $(Z^*)^2 = (Z^\circ)^\perp$ (true in general, cf. Lemma 5.1) and $Z^* = \delta(Z)^{-1}Z$ for any line Z (a speciality in dimension 2).

(ii) By (i), if $y \in (W^2)^\perp$ then (17) has 7 solutions (one for each Z). Since $Z \mapsto \delta(Z)$ is one-to-one, at most 3 solutions can be in W , i.e., correspond to $\delta(W)/\delta(Z)^3 = y$. Hence y is a value of multiplicity ≥ 4 in this case. On the other hand, if $y \notin (W^2)^\perp$ then (17) has a solution for one particular Z , which may be in W , and hence y has multiplicity 0 or 1. \square

Definition 6.3. The space $C = (W^2)^\perp \subset \mathbb{F}_{2^n}$ is called the *collision space* (of the RRP) relative to W . The matrix $\mathbf{C}_W = (c_{ij})$ whose rows are labelled with the solids of $\text{PG}(\mathbb{F}_{2^n})$ containing W , whose columns are labelled with the elements of $(W^2)^\perp$ (relative to some orderings of these sets), and whose (i, j) entry c_{ij} is defined

⁴⁹Orthogonality is taken with respect to the trace bilinear form.

as the number of planes $E \neq W$ in the solid T_i satisfying $\sigma_W(E) = y_j$, is called a *collision matrix* relative to W .⁵⁰

Since both \mathbb{F}_{2^n}/W and $(W^2)^\perp$ have dimension $n - 3$, \mathbf{C}_W is a square matrix of order $2^{n-3} - 1 = 2^{v-6} - 1$. From the preceding development it should be clear that \mathbf{C}_W contains the necessary information to determine the maximum net gain for all subsets $\mathcal{A} \subset \mathcal{G}_W$ satisfying the conditions of Lemma 4.1 and thus essentially solve the RRP. This is made explicit in our next theorem. The actual solution also requires selecting planes E with $\sigma_W(E) = y$ in the solids T corresponding to \mathcal{A} and finding the corresponding new planes in \mathcal{N}_1 , but this is a straightforward computational task and will not be discussed further.

Theorem 6.4. *Let W be a plane in $\text{PG}(\mathbb{F}_{2^n})$, $m = 2^{n-3} - 1$ the order of the corresponding collision matrix \mathbf{C}_W and r_i the i -th row sum of \mathbf{C}_W ($1 \leq i \leq m$). The maximum local net gain achievable in the RRP specialized to W is the solution of the following combinatorial optimization problem:*

$$(18) \quad \begin{array}{ll} \text{Maximize} & \sum_{i=1}^m (6 - r_i)x_i + \text{wHam}(\mathbf{x}\mathbf{C}_W) \\ \text{subject to} & \mathbf{x} \in \{0, 1\}^m \end{array}$$

As a consequence of this theorem and Corollary 2, the maximum local net gain achievable in the general RRP (which depends only on $n = v - 3$) is equal to the largest optimal solution of the family of optimization problems (18), with W running through a system of representatives for the G -orbits on planes in $\text{PG}(\mathbb{F}_{2^n})$.

Proof of Theorem 6.4. An admissible selection of $\mathcal{A} \subset \mathcal{G}_W$ corresponds to a set T_i , $i \in I \subseteq \{1, \dots, m\}$, of solids containing W (the indexing is the same as in the definition of \mathbf{C}_W) and hence to a unique vector $\mathbf{x} \in \{0, 1\}^m$ (the characteristic vector of I). Let $\mathbf{C}_W(I)$ be the submatrix of \mathbf{C}_W with rows indexed by $i \in I$, and $t = \#I$.

The number of planes E involved in the rearrangement (equal to $\#\mathcal{N}_1$) is $14t$, of which $\sum_{i \in I, 1 \leq j \leq m} c_{ij}$ have $\sigma_W(E) \in (W^2)^\perp$. The corresponding local net gain is obtained by selecting from the $14t$ planes all those which have $\sigma_W(E) \notin (W^2)^\perp$, and for each value $y \in (W^2)^\perp$ that corresponds to a nonzero column of $\mathbf{C}_W(I)$ one further plane with $\sigma(E) = y$. Hence the local net gain equals

$$14t - \sum_{i \in I, 1 \leq j \leq m} c_{ij} + \text{wHam}(\mathbf{x}\mathbf{C}_W) - 8t = \sum_{i \in I} (6 - r_i) + \text{wHam}(\mathbf{x}\mathbf{C}_W),$$

as claimed. \square

Before deriving further properties of collision matrices, we illustrate the newly developed concepts with examples, including a solution of the RRP in the case $v = 9$.

Example 5 ($v = 7$, continuation of Example 2). This case is now rather trivial. Recalling that $W \subset \mathbb{F}_{16}$ has been chosen as the trace-zero subspace, we only observe the following: The collision space in this case is $(W^2)^\perp = W^\perp = \mathbb{F}_2$; i.e., $\sigma(E) = 1$ is the only multiple value of the σ -invariant and at the same time a missing point of multiplicity 3. The corresponding collision matrix is the 1×1 matrix $\mathbf{C}_W = (4)$,

⁵⁰In what follows, we will often say “the collision matrix \mathbf{C}_W ”. This is slightly inaccurate but forgivable in our case, since collision matrices for the same W differ only by row and column permutations and all properties discussed will be invariant under these.

and (18) has objective value 3 (attained at $x = 1$). The 4 missing points outside the collision space are the primitive 5th roots of unity in \mathbb{F}_{16} .

Example 6 ($v = 8$, continuation of Example 3). In the case $v = 8$, $W = \{1, \alpha, \alpha^2\}$, the collision space is the line $\{\alpha^{23}, \alpha^{25}, \alpha^{28}\}$. The points of the trace zero subspace of \mathbb{F}_{32} are α^i , $i \in \{0, 1, 2, 4, 7, 8, 14, 15, 16, 19, 23, 25, 27, 28, 29, 30\}$, and $\text{Tr}(\alpha^{2s+t}) = 0$ for $s = 0, 1, 2$ and $t = 23, 25, 28$. The collision matrix in this case is

$$\mathbf{C}_W = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix},$$

and (18) reduces to a trivial optimization problem with maximum objective value 3 attained at all nonzero vectors $\mathbf{x} \in \{0, 1\}^3$.

Further, using $1 + \alpha^2 = \alpha^5$, $1 + \alpha = \alpha^{18}$, $\alpha + \alpha^2 = \alpha^{19}$, $1 + \alpha + \alpha^2 = \alpha^{11}$, we have $W = \{1, \alpha, \alpha^2, \alpha^5, \alpha^{11}, \alpha^{18}, \alpha^{19}\}$, $\delta(W) = \alpha^{25}$; the lines in W are $L_1 = \{1, \alpha, \alpha^{18}\}$, $L_2 = \{1, \alpha^2, \alpha^5\}$, $L_3 = \{1, \alpha^{11}, \alpha^{19}\}$, $L_4 = \{\alpha, \alpha^2, \alpha^{19}\}$, $L_5 = \{\alpha, \alpha^5, \alpha^{11}\}$, $L_6 = \{\alpha^{18}, \alpha^2, \alpha^{11}\}$, $L_7 = \{\alpha^{18}, \alpha^5, \alpha^{19}\}$, and $W' = \{\delta(Z); Z \subset W\} = \{1, \alpha^7, \alpha^{11}, \alpha^{17}, \alpha^{19}, \alpha^{22}, \alpha^{30}\}$. Hence the missing points are $\{\delta(W)/\delta(Z)^3; Z \subset W\} = \{\alpha^4, \alpha^5, \alpha^{21}, \alpha^{23}, \alpha^{25}, \alpha^{28}, \alpha^{30}\}$. Note that all points on the collision line are missing points of multiplicity 1. We will see later (cf. Theorem 6.5(i)) that this fact is responsible for the three 2's in each column of \mathbf{C}_W .

Example 7 ($v = 9$). Here $n = 6$ and the corresponding extension field is $\mathbb{F}_{64} = \mathbb{F}_2[\alpha]$ with $\alpha^6 + \alpha^4 + \alpha^3 + \alpha + 1 = 0$. The $\begin{bmatrix} 6 \\ 3 \end{bmatrix}_2 = 1395$ planes in $\text{PG}(\mathbb{F}_{64})$ fall into 7 G -orbits with representatives $W_1 = \langle 1, \alpha, \alpha^2 \rangle$, $W_2 = \langle 1, \alpha, \alpha^3 \rangle$, $W_3 = \langle 1, \alpha, \alpha^4 \rangle$, $W_4 = \langle 1, \alpha, \alpha^5 \rangle$, $W_5 = \langle 1, \alpha, \alpha^{22} \rangle$, $W_6 = \langle 1, \alpha^3, \alpha^{18} \rangle$, $W_7 = \langle 1, \alpha^9, \alpha^{18} \rangle$ and orbit lengths 189, 378, 126, 378, 189, 126, 9, respectively. The corresponding collision matrices are, in order,

$$\begin{pmatrix} 1 & 1 & 2 & 1 & 2 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 2 & 1 & 2 & 0 & 1 \\ 1 & 1 & 2 & 1 & 2 & 4 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 2 \\ 2 & 2 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & 1 & 1 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 & 1 & 1 & 2 \\ 0 & 2 & 1 & 1 & 1 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 2 & 1 & 1 & 2 & 1 & 1 \\ 1 & 2 & 1 & 1 & 0 & 1 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 2 & 0 & 0 & 1 & 2 & 1 & 0 \\ 2 & 0 & 2 & 1 & 0 & 1 & 2 \\ 2 & 2 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 2 \\ 0 & 2 & 2 & 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 2 & 1 & 2 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 & 1 & 0 & 2 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 & 1 & 0 & 2 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 & 1 & 0 & 2 \\ 1 & 1 & 0 & 0 & 1 & 4 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 0 & 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 & 2 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 & 2 \\ 2 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 0 & 2 & 2 & 0 \\ 2 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 2 & 0 & 2 \end{pmatrix}.$$

The computations were done with SageMath.

From this point onward it is fairly easy to solve the RRP by hand. A closer look at (18) reveals that rows i with $r_i < 6$ must be part of any optimal solution (since they strictly increase the net gain) and those with $r_i = 6$ can be included w.l.o.g. in any optimal solution (since they cannot decrease the net gain). Moreover, since $w_{\text{Ham}}(\mathbf{x}\mathbf{C}_W) \leq m$, the optimal local net gain is upper bounded by $\sum_{i; r_i < 6} (6 - r_i) + m$.

These observations give that the 1st and 6th collision matrix have a maximum local net gain ≥ 12 (corresponding to $I = \{2, 3, 4, 5\}$ and $I = \{1, 3, 5\}$, respectively) and allows us to discard the other 5 collision matrices, whose optimal values are

bounded by 9, 7, 9, 9, and 7 (in that order). Then it is easy to complete the solution: The overall maximum local net gain is 12, and is attained precisely for the following \mathbf{C}_W and I : $I = \{2, 3, 4, 5\}$, $\{1, 2, 3, 4, 5\}$, $\{2, 3, 4, 5, 6\}$ for the 1st collision matrix (corresponding to $W = \langle 1, \alpha, \alpha^2 \rangle$) and $I = \{1, 3, 5\}$, $\{1, 3, 5, 7\}$, $\{1, 2, 3, 5\}$, $\{1, 3, 4, 5\}$, $\{1, 3, 5, 6\}$, $\{1, 2, 3, 5, 7\}$, $\{1, 3, 4, 5, 7\}$, $\{1, 3, 5, 6, 7\}$, for the 6th collision matrix (corresponding to $W = \langle 1, \alpha^3, \alpha^{18} \rangle$).

Thus the largest subspace codes that are obtained as solutions of the RRP for $v = 9$ have size $2^{12} + 12 \cdot 63 = 4852$ and exceed the LMRD code bound $2^{12} + \begin{bmatrix} 6 \\ 3 \end{bmatrix}_2 = 4747$.

Finally note that the collision matrices with optimal value 12 are exactly those which have an entry $c_{ij} = 4$.

Now we examine the collision matrices \mathbf{C}_W in more detail. Clearly \mathbf{C}_W is non-negative and integer-valued, and it appears from the preceding examples that the only values which occur in \mathbf{C}_W are 0, 1, 2, 4 and the distribution of these values in every column is restricted to a few different types. In our next theorem we will prove this and several other properties of \mathbf{C}_W , which facilitate the solution of the optimization problem (18). In the statement of the theorem we use the *type* (or *spectrum*) of a row or column, which refers to the multiset of its entries, with 0 omitted. Thus, e.g., $1^4 2^2$ refers to a row or column of \mathbf{C}_W containing four 1's, two 2's and $2^{n-3} - 1 - 6$ zeros.

Theorem 6.5. *Let W be a plane in $\text{PG}(\mathbb{F}_{2^n})$ and $\mathbf{C}_W \in \mathbb{Z}^{m \times m}$, $m = 2^{n-3} - 1$, the associated collision matrix.*

- (i) *The columns of \mathbf{C}_W have type 1^7 , 2^3 , or 4^1 . More precisely, a column labeled with $y \in (W^2)^\perp$ has type 1^7 if y is not a missing value of σ_W (i.e., $y \neq \delta(W)/\delta(Z)^3$ for all lines $Z \subset W$), type 2^3 if y is a missing value of multiplicity 1 (i.e., $y = \delta(W)/\delta(Z)^3$ for exactly one line $Z \subset W$), and type 4^1 if y is a missing value of multiplicity 3 (i.e., $y = \delta(W)/\delta(Z)^3$ for three lines $Z \subset W$). Moreover, Type 4^1 does not occur if n is odd, and occurs at most once as a column of \mathbf{C}_W if n is even.*
- (ii) *The support of each column forms a subspace of \mathbb{F}_{2^n}/W (a plane if the type is 1^7 , a line if the type is 2^3 and, trivially, a point if the type is 4^1).*
- (iii) *All rows of \mathbf{C}_W have the same parity, equal to the parity of the number of columns of type 1^7 .⁵¹*

Proof. (i) First we show that the multiplicities of y as a missing point of σ_W and their occurrences must be as indicated. The maximum multiplicity is 3, since $\delta(Z)^3 = \delta(W)/y$ can have at most 3 solutions Z (cf. Lemma 3.3 and Theorem 5.2(ii)). If there are two different solutions Z_1, Z_2 then $\omega = \delta(Z_2)/\delta(Z_1)$ must be a primitive 3rd root of unity in \mathbb{F}_{2^n} , which forces $n \equiv 0 \pmod{2}$. Moreover, denoting the third line in W through the intersection point $Z_1 \cap Z_2$ by Z_3 , we then have

$$\omega^2 \delta(Z_1) = \delta(Z_1) + \delta(Z_2) = \delta(Z_3),$$

and Z_3 is a third solution. Since the line $\{\delta(Z_1), \delta(Z_2), \delta(Z_3)\} = \mathbb{F}_4^\times \delta(Z_1)$ is a member of the standard line spread of $\text{PG}(\mathbb{F}_{2^n})$, the plane $W' = \{\delta(Z); Z \subset W\}$ cannot contain a further such line, showing that there is at most one missing value of multiplicity 3.

⁵¹This property may seem trivial from the shape of the collision matrices in Example 7, but for $v > 9$ there are no all-one columns and hence this property is no longer obvious.

Next we set $\{Z; Z \subset W\} = \{Z_i; 1 \leq i \leq 7\}$ and let $E_i \supset Z_i$ be the corresponding plane satisfying $\delta(E_i)/\delta(Z_i)^3 = y$ ($E_i = W$ is allowed here). Using the alternative expression for $\delta(E_i)$ in terms of $\delta(Z_i)$ and $s_{Z_i}(X)$, we can write these equations as $s_{Z_i}(x_i)/\delta(Z_i)^2 = y$, $x_i \in E_i \setminus Z_i$. Using (12), we obtain

$$\begin{aligned} s_W(x_i) &= s_{Z_i}(x_i)^2 + s_{Z_i}(c_i)s_{Z_i}(x_i) \\ &= (y\delta(Z_i)^2)^2 + s_{Z_i}(c_i)y\delta(Z_i)^2 \\ &= y^2\delta(Z_i)^4 + y\delta(W)\delta(Z_i) \end{aligned}$$

with $c_i \in W \setminus Z_i$. This shows that $s_W(x_i) = f(\delta(Z_i))$ is in the image of the plane $W' = \{\delta(Z); Z \subset W\} = \delta(W)W^*$ under the \mathbb{F}_2 -linear transformation $f(x) = y^2x^4 + y\delta(W)x$. But $\text{Ker}(f|_{W'})$ consists of 0 and all elements $\delta(Z_i)$ satisfying $\delta(W)/\delta(Z_i)^3 = y$, and hence has dimension 0, 1, or 2. Applying the homomorphism theorem for linear maps, the remaining assertions of (i) follow.⁵²

(ii) This has been already shown as part of the proof of (i).

(iii) If $(W^2)^\perp$ contains μ missing points of σ_W , the number of columns of \mathbf{C}_W of type 1^7 is equal to $2^{n-3} - 1 - \mu$.

On the other hand, consider a solid $T_i \supset W$. For any line $Z \subset W$, there are 2 planes $E_1, E_2 \subset T_i$ such that $E_1 \cap W = E_2 \cap W = Z$. The image of $\{E_1, E_2, W\}$ under $E \mapsto \delta(E)/\delta(Z)^3$ is a line in $\text{PG}(\mathbb{F}_{2^n})$ through the missing point $\delta(W)/\delta(Z)^3$. Hence the restriction of σ_W to the 14 planes $E \neq W$ in T_i determines 7 lines, one line through each missing point,⁵³ containing the 14 values $\sigma_W(E)$. Since $(W^2)^\perp$ forms a hyperplane in the image $\delta(Z)^{-2}Z^\circ$, these lines are either contained in $(W^2)^\perp$ or meet $(W^2)^\perp$ in a unique point. Hence Z contributes 0 or 2 to the row sum r_i if $\delta(W)/\delta(Z)^3 \in (W^2)^\perp$, and 1 to r_i if $\delta(W)/\delta(Z)^3 \notin (W^2)^\perp$.⁵⁴ The parity of r_i is thus equal to $7 - \mu$. But $7 - \mu \equiv 2^{n-3} - 1 - \mu \pmod{2}$, and the proof is complete. \square

As we have seen in Example 7, knowledge of the number of rows of \mathbf{C}_W with $r_i \leq 6$ provides important information about the optimal solutions of (18) and, in particular, can be used to bound the maximum net gain achievable when using W . In view of its importance, we now state this result in the general case. The *row-sum spectrum* of \mathbf{C}_W refers to the multiset of row sums of \mathbf{C}_W and is denoted by $0^{m_0}1^{m_1}2^{m_2} \dots$ if there are m_r rows with row sum r .

Corollary 4. *Suppose \mathbf{C}_W has row-sum spectrum $0^{m_0}1^{m_1}2^{m_2} \dots$, and the union of the supports of the $m_0 + m_1 + \dots + m_6$ rows of \mathbf{C}_W with $r_i \leq 6$ (equal to the number of nonzero columns of the corresponding submatrix $\mathbf{C}_W(I)$) is lower-bounded by m' . Then the optimal value N_1 of (18) (i.e., the maximum net gain achievable relative to W) satisfies the bounds*

$$\sum_{r=0}^5 m_r(6-r) + m' \leq N_1 \leq \sum_{r=0}^5 m_r(6-r) + m,$$

⁵²It should be noted that $s_W(x_i) = y^2\delta(Z_i)^4 + y\delta(W)\delta(Z_i)$ is equivalent to $s_{Z_i}(x_i)/\delta(Z_i)^2 = y \vee s_{Z_i}(x_i + c_i)/\delta(Z_i)^2 = y$. Both planes $\langle Z_i, x_i \rangle$, $\langle Z_i, x_i + c_i \rangle$ are in the same solid $T = \langle W, x_i \rangle$ and only one of them can be a solution of the equation. The number of solutions in any “point” T of \mathbb{F}_{2^n}/W is thus invariant under the transformation just made.

⁵³Again the missing points are counted with their multiplicity.

⁵⁴In the first case, the contribution is 0 if the corresponding line meets $(W^2)^\perp$ in the missing point $\delta(W)/\delta(Z)^3$, and 2 otherwise.

where $m = 2^{n-3} - 1$ is the order of \mathbf{C}_W .

Proof. This is immediate from Theorem 6.4. \square

As we will see in a moment, the row-sum spectrum of \mathbf{C}_W depends only on the geometric configuration of the (at most 7) missing points contained in $(W^2)^\perp$ and hence is quite restricted. Finding a good lower bound m' without actually computing \mathbf{C}_W seems to be more difficult. A reasonable approach to solve this problem is to find a good upper bound l for the column sums of $\mathbf{C}_W(I)$ and use the obvious fact that the number of nonzero columns of $\mathbf{C}_W(I)$ must be at least $(\sum_{r=0}^6 rm_r)/l$.⁵⁵

From Theorem 6.2 we have that for every line $Z \subset W$ there exists a unique hyperplane $H_Z \supset Z$ which is mapped onto $(W^2)^\perp$ by $x \mapsto \delta(\langle Z, x \rangle)/\delta(Z)^3 = s_Z(x)/\delta(Z)^2$.

- Theorem 6.6.** (i) *The hyperplane H_Z has equation $\text{Tr}(\frac{\delta(W)}{\delta(Z)^3} \cdot x^2) = 0$ and hence is essentially the dual of the corresponding missing point of σ_W under the trace bilinear form;*
- (ii) *$H_Z \supset W$ iff $\frac{\delta(W)}{\delta(Z)^3} \in (W^2)^\perp$ iff $\text{Tr}(a_1 c^4/a_0^2) = 0$, where $a_1 = \delta_1(Z)$, $a_0 = \delta(Z)$ and $c \in W \setminus Z$.*
- (iii) *For any solid $T_i \supset W$ the number of planes $E \neq W$ contained in T_i and satisfying $\sigma_W(E) \in (W^2)^\perp$ (i.e., the row sum r_i of \mathbf{C}_W) is equal to $7 - \mu + 2\nu$, where μ denotes the number of missing points of σ_W contained in $(W^2)^\perp$ and ν the number of hyperplanes H_Z that contain T_i .*

Note that H_Z can contain T_i only if it contains W . Hence the ν hyperplanes in (iii) are among those μ with their corresponding missing point in $(W^2)^\perp$, and we can restate the formula $r_i = 7 - \mu + 2\nu$ in the following way: A hyperplane H_Z contributes 0, 1, or 2 to the row sum r_i if $H_Z \supset W$ and $H_Z \not\supseteq T_i$, $H_Z \not\supseteq W$, or $H_Z \supset T_i$, respectively. For conditions equivalent to $H_Z \supset W$ see (ii).

Proof of Theorem 6.6. (i) $x \in H_Z$ is equivalent to

$$\langle x, y \rangle_Z = \text{Tr} \left(\frac{s_Z(x)y^2}{\delta(Z)^2} \right) = 0 \quad \text{for all } y \in W.$$

Since $\langle \cdot, \cdot \rangle_Z$ is symmetric and has radical Z , this is equivalent to $\langle x, c \rangle_Z = \langle c, x \rangle_Z = 0$ for any $c \in W \setminus Z$, i.e. to $\text{Tr} \left(\frac{s_Z(c)x^2}{\delta(Z)^2} \right) = \text{Tr} \left(\frac{\delta(W)x^2}{\delta(Z)^3} \right) = 0$.

(ii). As in (i), $H_Z \supset W$ is equivalent to $\langle c, c \rangle_Z = \text{Tr} \left(\frac{s_Z(c)c^2}{\delta(Z)^2} \right) = 0$, which in turn is equivalent to $\langle c, y \rangle_Z = 0$ for all $y \in W$ and hence to $\frac{\delta(W)}{\delta(Z)^3} = \frac{s_Z(c)}{\delta(Z)^2} \in (W^2)^\perp$. The second equivalence follows from $\langle c, c \rangle_Z = \text{Tr}(a_1 c^4/a_0^2)$; cf. Lemma 6.1.

(iii) is proved using (ii) and the reasoning in the proof of Theorem 6.5(iii). The case $\frac{\delta(W)}{\delta(Z)^3} \in (W^2)^\perp$ now splits into two subcases according to whether the image of $\{E_1, E_2, W\}$ is contained in $(W^2)^\perp$ or not. The first case is equivalent to $T_i \subseteq H_Z$ and accounts for 2 values $\sigma(E_1) = \sigma(E_2) \in (W^2)^\perp$, the second case for 0 values. \square

⁵⁵The obvious bound $l \leq 7$ won't do the job, of course, since the row sums of $\mathbf{C}_W(I)$ are ≤ 6 and hence no constant $l \geq 6$ can improve on the trivial bound m .

Using Theorem 6.6, the row-sum spectrum of \mathbf{C}_W can be determined from the multiset \mathbf{m} of missing points contained in $(W^2)^\perp$ in the same way as the weight enumerator of a binary linear $[\mu, k]$ code with associated multiset \mathbf{m} , represented by the columns of a generator matrix of the code. For the latter it is usually assumed that the multiset spans the geometry, which in our case need not be true. However, it is easy to reduce the spectrum computation to this case: Denoting by M the hull of \mathbf{m} (i.e., the subspace generated by the missing points in $(W^2)^\perp$), we compute the associated weight distribution $(A_i)_{0 \leq i \leq \mu}$, replace nonzero weights i by the corresponding row sums $2(\mu - i) + 7 - \mu = \mu + 7 - 2i$ and scale the frequencies A_i by $2^{n-3-\dim(M)}$. If M is a proper subspace of $(W^2)^\perp$, there are in addition $2^{n-3-\dim(M)}$ rows of \mathbf{C}_W corresponding to the all-zero codeword. These correspond to the solids T_i contained in $(M^{1/2})^\perp = \bigcap \{H_Z; H_Z \supset W\}$ and have maximum row sum $r_i = 7 - \mu + 2\mu = 7 + \mu$.⁵⁶ We will illustrate row-sum spectrum computations later in the proofs of Theorems 8.1 and 8.2.

For all even n (i.e., odd packet lengths v) explored so far, the maximum net gain of the RRP is achieved only by planes W whose collision matrices have an entry $c_{ij} = 4$. It is therefore of interest, to characterize these planes. For the statement of the following theorem, we denote the trace-zero hyperplane of $\text{PG}(\mathbb{F}_{2^n})$ by H_2 .⁵⁷

Theorem 6.7. (i) Suppose n is even and ω is a generator of the subfield $\mathbb{F}_4 \subset \mathbb{F}_{2^n}$. A plane W in $\text{PG}(\mathbb{F}_{2^n})$ gives rise to an entry $c_{ij} = 4$ in the collision matrix \mathbf{C}_W if and only if $W = rW_1$, $r \in \mathbb{F}_{2^n}^\times$, for some plane $W_1 = \langle 1, a, b \rangle$ with a, b satisfying $b^2 + b = \omega(a^2 + a)$.
(ii) The planes W_1 of the type indicated in (i) are contained in H_2 , mutually intersect in the point $\mathbb{F}_2 = \mathbb{F}_2 1$ of $\text{PG}(\mathbb{F}_{2^n})$, and determine a line spread of the “sub-quotient” geometry $\text{PG}(H_2/\mathbb{F}_2) \cong \text{PG}(n-3, \mathbb{F}_2)$. In particular, the number of such planes is $(2^{n-2} - 1)/3$.
(iii) For a plane W_1 of the type indicated in (i), the missing points of σ_{W_1} are 1 (of multiplicity 3) and $(b + \omega a + x)^{-3}$ for $x \in \mathbb{F}_4$ (of multiplicity 1).

Proof. (i) Since the indicated property is G -invariant, we may assume $1 \in W$ and that the three lines Z_1, Z_2, Z_3 containing 1 give rise to the missing point of multiplicity 3, i.e. $\delta(Z_1)^3 = \delta(Z_2)^3 = \delta(Z_3)^3$.

Now let $Z_1 = \langle 1, a \rangle$, $Z_2 = \langle 1, b \rangle$, and hence $Z_3 = \langle 1, a + b \rangle$. Then $b^2 + b = \delta(Z_2) = \omega^i \delta(Z_1) = \omega^i(a^2 + a)$ for some $i \in \{1, 2\}$, and by interchanging a, b , if necessary, we may assume $i = 1$.

Conversely, assume that $W = \langle 1, a, b \rangle$ with a, b having the indicated property. Then, with Z_i as in (i), we have $\delta(Z_i)^3 = (a^2 + a)^3$, $\delta(W) = \delta(Z_1)\delta(Z_2)\delta(Z_3)/1^2 = \omega^{0+1+2}(a^2 + a)^3 = (a^2 + a)^3$, and hence the triple missing point is $\delta(W)/\delta(Z_i)^3 = 1$. Further, since $\text{Tr}_{\mathbb{F}_{2^n}/\mathbb{F}_4}(a^2 + a) = \text{Tr}(a)$ and similarly for b , we must have $\text{Tr}(b) = \omega \text{Tr}(a)$ and hence $\text{Tr}(a) = \text{Tr}(b) = 0$. This implies $1 \in (W^2)^\perp$, and hence $y = 1$ gives rise to a column of \mathbf{C}_W of Type 4¹ by Theorem 6.5(i).⁵⁸

⁵⁶All other row sums are $\leq 7 - \mu + 2(\mu - 1) = 5 + \mu$.

⁵⁷Thus $H_2 = \{x \in \mathbb{F}_{2^n}; \text{Tr}_2(x) = 0\}$, where $\text{Tr}_2(x) = \text{Tr}(x) = x + x^2 + x^4 + \dots + x^{2^{n-1}}$. The index used is thus equal to the order of base field of the corresponding field extension.

⁵⁸A more geometric proof of the fact that a missing point of multiplicity 3 must be in $(W^2)^\perp$ is the following: Consider the 7 lines determined by the restriction of σ_W to a fixed solid $T_i \supset W$; cf. the proof of Theorem 6.5(iii). Since these lines are contained in the corresponding spaces $\delta(Z)^{-2}Z^\circ$, they can only intersect in $(W^2)^\perp$. However, the 3 lines containing the triple missing point intersect in this point, and hence this point must be in $(W^2)^\perp$.

(ii) In the proof of (i) we have seen that such planes W_1 are contained in H_2 . Since $\text{Tr}(a) = 0$ is equivalent to $\text{Tr}_{\mathbb{F}_{2^n}/\mathbb{F}_4}(a^2 + a) = 0$, the map $\mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^n}$, $x \mapsto x^2 + x$ induces an isomorphism from H_2/\mathbb{F}_2 onto the trace-zero subspace H_4 of the field extension $\mathbb{F}_{2^n}/\mathbb{F}_4$. By Part (i), it maps the set of planes W_1 of the indicated type onto the standard line spread in $\text{PG}(\mathbb{F}_{2^n}/\mathbb{F}_4)$. The result follows.

(iii) We know already that 1 is the missing point of multiplicity 3. Since the plane $\{\delta(Z); Z \subset W_1 \text{ a line}\}$ is generated by $\mathbb{F}_4(a^2 + a)$ and $\delta(\langle a, b \rangle) = ab^2 + a^2b$, the remaining 4 missing points are

$$\frac{(a^2 + a)^3}{(ab^2 + a^2b + x(a^2 + a))^3} = \frac{(a^2 + a)^3}{(ab + \omega(a^3 + a^2) + x(a^2 + a))^3} = \frac{1}{(b + \omega a + x)^3}$$

with $x \in \mathbb{F}_4$, as asserted. \square

Remark 4. The map $\mathbb{F}_{2^n} \rightarrow \mathbb{F}_{2^n}$, $x \mapsto x^2 + x$ induces also an \mathbb{F}_2 -isomorphism from $\mathbb{F}_{2^n}/\mathbb{F}_4$ onto H_2/\mathbb{F}_2 , hence gives rise to the chain $\mathbb{F}_{2^n}/\mathbb{F}_4 \rightarrow H_2/\mathbb{F}_2 \rightarrow H_4$ of \mathbb{F}_2 -isomorphisms.⁵⁹

The points $\mathbb{F}_4(b + \omega a)$ with $a \in H_2 \setminus \mathbb{F}_2$ and b determined as in Theorem 6.7(i) form a system of representatives for the nonzero cosets in $\mathbb{F}_{2^n}/\mathbb{F}_4$ and hence for the lines in $\text{PG}(\mathbb{F}_{2^n}/\mathbb{F}_4) \cong \text{PG}(n/2 - 1, \mathbb{F}_4)$ that pass through the point $\mathbb{F}_4 = \mathbb{F}_4 1$. This can be seen as follows: Since

$$\begin{aligned} (b + \omega a)^2 + b + \omega a &= b^2 + b + \omega^2 a^2 + \omega a = a^2, \\ (\omega b + \omega^2 a)^2 + \omega b + \omega^2 a &= \omega^2 b^2 + \omega b + \omega a^2 + \omega^2 a = b^2 + a^2, \end{aligned}$$

the line $L = \mathbb{F}_4(b + \omega a) + \mathbb{F}_4$ is mapped to the plane $W_1^2 = \langle 1, a^2, b^2 \rangle$ by $x \mapsto x^2 + x$, and the planes of this form partition H_2/\mathbb{F}_2 ; cf. Theorem 6.7(ii).

Moreover, by Theorem 6.7(iii) the missing points of σ_{W_1} are just the reciprocal cubes of the 5 points on the line $L = \mathbb{F}_4(b + \omega a) + \mathbb{F}_4$.⁶⁰

These observations imply that each element $\neq 1$ of the index-3 subgroup of $\mathbb{F}_{2^n}^\times$ forms a missing point for precisely one plane W_1 of the type indicated in Theorem 6.7(i).

As an aside, making the link with Section 5, we note that the plane polynomials of the planes $W_1 = \langle 1, a, b \rangle$ in Theorem 6.7 are $s_{W_1}(X) = (X^4 + (a^2 + a)^3 X) \circ (X^2 + X) = X^8 + X^4 + (a^2 + a)^3 X^2 + (a^2 + a)^3 X$.

7. COMPUTATIONAL RESULTS

In this section we provide an account of explicit maximum net gain computations for $10 \leq v \leq 15$, which we have done using the computer algebra package SageMath. The computations were exhaustive for $v \leq 13$. In the case $v = 13$ ($n = 10$) there are 633 G -orbits to process. For each G -orbit representative W we have computed the collision matrix \mathbf{C}_W (of size 127×127 for $v = 13$) and the bounds for the maximum net gain N_1 relative to W stated in Corollary 4. Then, in a second pass through the list of G -orbit representatives, this time sorted in order of decreasing lower bounds for N_1 , we have computed the exact maximum net gains N_1 for those G -orbits, for which the upper bound still exceeded the current “absolute” maximum net gain (taken over all G -orbits computed so far). The actual optimization routine used

⁵⁹This property is reflected in the symbolic factorization $X^4 + X = (X^2 + X) \circ (X^2 + X)$.

⁶⁰Note that the cube x^3 of a point $\mathbb{F}_4 x$ is well-defined.

some greedy heuristic for selecting rows of \mathbf{C}_W with row sums > 6 as part of the next-to-be-tested feasible solution.

For $v \in \{14, 15, 16\}$ exhaustive computations were not feasible, and we have restricted the search to those G -orbits, which contain a plane W_1 of the type discussed in Theorem 6.7, or a subset thereof. In Section 9 we will show that the absolute maximum net gains obtained for $v \in \{14, 15\}$ nevertheless represent the true maximum as well.

The computational results are summarized in Table 1, including the cases $7 \leq v \leq 9$ already discussed. The table contains for each length v the number of G -orbits processed (for $v \geq 13$ equal to the total number of G -orbits), the absolute maximum local net gain N_1 computed (with the possible exception of $v = 16$ equal to the true maximum), the local net gain equivalent of the LMRD code bound (“LMRD threshold”), the size of the plane subspace codes corresponding to optimal solutions of the RRP, and a representative subspace W giving rise to an (absolutely) optimal solution. The generators of W are given as powers of a primitive α of \mathbb{F}_{2^n} (root of the Conway polynomial of degree n , as used by SageMath). The next few paragraphs contain supplementary remarks on each case.

$v = 7$. According to [36, 29], there exist solutions \mathcal{C} of the RRP that can be extended by 28 further planes meeting $S = \{0\} \times \mathbb{F}_{16}$ in a line to a currently best known $(7, 329, 4; 3)_2$ code. However, no rotation-invariant $(7, 301, 4; 3)_2$ code \mathcal{C} has this property.⁶¹

$v = 8$. Using a modified beam-search algorithm [11], we have found that one of the optimal $(8, 1117, 4; 3)_2$ solutions \mathcal{C} of the RRP can be augmented by 142 extra planes meeting $S = \{0\} \times \mathbb{F}_{32}$ in a line to a $(8, 1259, 4; 3)_2$ code. This is considerably better than the LMRD code bound $1024 + 155 = 1179$, but it falls short of the currently best known code of size 1326.

$v = 9$. The seven 7×7 collision matrices corresponding to the seven G -orbits were already listed in Example 7. Two G -orbits, with orbit representatives $\langle 1, \alpha, \alpha^2 \rangle$ and $\langle 1, \alpha^3, \alpha^{18} \rangle$, yield the absolute maximum local net gain 12, resulting in $(9, 4852, 4; 3)_2$ codes. We have found that 162 planes meeting $S = \{0\} \times \mathbb{F}_{64}$ can be added to one of the codes, increasing the code size to 5014. The currently best known code has size 5986 [10].

$v = 10$. In this case all fifteen 15×15 collision matrices were computed. The absolute maximum local net gain 20 is obtained from the three G -orbits with representatives $\langle 1, \alpha, \alpha^{24} \rangle$, $\langle 1, \alpha, \alpha^{39} \rangle$ and $\langle 1, \alpha, \alpha^{22} \rangle$, resulting in $(10, 18924, 4; 3)_2$ codes. The size of these codes is smaller than the LMRD code bound $2^{14} + \left[\frac{7}{2} \right]_2 + 19051$, but again a further extension step by planes meeting $\{0\} \times \mathbb{F}_{128}$ in a line (in this case 1593 codewords can be added to one of the codes) increases the code size to $20517 > 19051$. The currently best known code has size 23870 [10].

$v = 11$. Here we have $n = 8$ and the collision matrices have already size 31×31 . Among the 53 G -orbits, the orbit containing $W = \langle 1, \alpha^{17}, \alpha^{34} \rangle = \{x \in \mathbb{F}_{16}; \text{Tr}(x) = 0\}$ uniquely gives the absolute maximal local net gain 54, resulting in a subspace code of size $2^{16} + 54 \cdot (2^8 - 1) = 79306$. This is better than the LMRD code bound $2^{16} + 10795 = 76331$, but should also be compared to the size 97526 of the currently best known code [10]. The collision matrix \mathbf{C}_W is shown in Figure 1. This case is particularly important, since it serves as the “anchor” case for the family of packet

⁶¹M. Kiermaier, personal communication

$$\begin{pmatrix} 00011002100110000000001100000010 \\ 0000011001100000000010011000012 \\ 02100000000000001111000000001110 \\ 00000000000000000101000000000000 \\ 000000000000001000000000000100000 \\ 00001000000010000000000000000000 \\ 00001000000010101010000000010000 \\ 00000110100000000000001111000012 \\ 40100000111000000000011100001110 \\ 001001100000000000000000011001112 \\ 0000000001100110000210000110010 \\ 00000000000000010000000000010000 \\ 00000000000000001010000000000000 \\ 000100000001010101010000000100000 \\ 00010000000100000000000000000000 \\ 00000000011000000000001100001100 \\ 00100000000000110000211100110000 \\ 00111002111110000000000000000000 \\ 000110021001100000000100000001100 \\ 00001100000001100000000001100000 \\ 00001010000010010100000010000000 \\ 000001000000001101010000001010000 \\ 00000010000000000000000010000000 \\ 021000000000000001111011100000000 \\ 000000000110000000000011000001100 \\ 000000001000001100002100000111100 \\ 02100000111000001111000000000000 \\ 0001010000001000001010000001000000 \\ 000100100000100100000000010010000 \\ 00000100000000000000000001000000 \\ 000000100000001001010000010100000 \end{pmatrix}$$

FIGURE 1. The collision matrix \mathbf{C}_W for $v = 11$ ($n = 8$) and $W = \langle 1, \alpha^{17}, \alpha^{34} \rangle = \{x \in \mathbb{F}_{16}; \text{Tr}(x) = 0\}$

lengths $v \equiv 3 \pmod{8}$ considered in Theorem 8.2 and is thus “responsible” for the constant $81/64$ in the bound in Part (ii) of our main theorem.

$v \geq 12$. In the cases $v = 12, 13$ we were still able to process all G -orbits and compute the absolute maximum local net gains directly; cf. Table 1. For lengths $v > 13$, however, the amount of calculation is too large for processing all G -orbits exhaustively. Hence in the remaining cases $v = 14, 15, 16$ we have processed only those G -orbits which appeared to be most “promising” in the sense that the lower bound in Corollary 4 is largest. The lower bound tends to be an increasing function of the number μ of missing points contained in the collision space $(W^2)^\perp$ and, in the case of odd v (even n) to be maximized for the planes W discussed in Theorem 6.7.

In the case $v = 14$ we processed all 513 G -orbits with $\mu \geq 5$ missing points in $(W^2)^\perp$. There are 381, 118, 14 G -orbits corresponding to $\mu = 5, 6, 7$, respectively, and the absolute maximum local net gain 379 is attained uniquely at a G -orbit with $\mu = 7$ (as was to be expected).

For $v = 15$, we processed all 34 G -orbits containing planes W as in Theorem 6.7. It turned out the absolute maximum local net gain 924 is attained at a particular G -orbit with $\mu = 3$, i.e., all 4 missing points of multiplicity 1 outside $(W^2)^\perp$.

Finally, for $v = 16$ we just processed all G -orbits with $\mu = 7$ and found for those an absolute maximum local net gain of 1526. This is better than the LMRD code bound, which is equivalent to a local net gain of 1365.

Thus it appears that $v = 8, 10$ are the only cases where the optimal solutions of the RRP have size smaller than the LMRD code bound; cf. also Conjecture 1 in Section 8.

v	n	$\#G\text{-orbits}$	N_1	$(N_1)_{\text{LMRD}}$	$\#\mathcal{C}$	W
7	4	1	3	2.33	$2^8 + 45$	$\langle 1, \alpha, \alpha^2 \rangle$
8	5	1	3	5.00	$2^{10} + 93$	$\langle 1, \alpha, \alpha^2 \rangle$
9	6	7	12	10.33	$2^{12} + 756$	$\langle 1, \alpha, \alpha^2 \rangle$
10	7	15	20	21.00	$2^{14} + 2540$	$\langle 1, \alpha, \alpha^{22} \rangle$
11	8	53	54	42.33	$2^{16} + 13770$	$\langle 1, \alpha^{17}, \alpha^{34} \rangle$
12	9	177	93	85.00	$2^{18} + 47523$	$\langle 1, \alpha^3, \alpha^{71} \rangle$
13	10	633	234	170.33	$2^{20} + 239382$	$\langle 1, \alpha, \alpha^{49} \rangle$
14	11	513	379	341.00	$2^{22} + 775813$	$\langle 1, \alpha^3, \alpha^{419} \rangle$
15	12	34	924	682.33	$2^{24} + 3783708$	$\langle 1, \alpha^{195}, \alpha^{1170} \rangle$
16	13	240	1526	1365.00	$2^{26} + 12499466$	$\langle 1, \alpha^{25}, \alpha^{1208} \rangle$

TABLE 1. Summary of maximum net gain computations

8. INFINITE FAMILIES OF SUBSPACE CODES EXCEEDING THE LMRD CODE BOUND

We are now in a position to compute explicit lower bounds for the maximum achievable net gain in the general RRP for packet lengths $v \equiv 3 \pmod{4}$ ($n \equiv 0 \pmod{4}$), using a careful choice for the plane W . It turns out that the corresponding modified subspace codes exceed the LMRD code bound. The analysis will be split into two cases depending on $v \pmod{8}$. We start with the easier case $v \equiv 7 \pmod{8}$.

Theorem 8.1. *For packet lengths $v \equiv 7 \pmod{8}$, i.e., $n = v - 3 \equiv 4 \pmod{8}$, the maximum achievable local net gain N_1 in the general RRP satisfies $N_1 \geq 3 \cdot 2^{n-4} = 3 \cdot 2^{v-7}$, and hence the corresponding optimum subspace codes have size*

$$\#\mathcal{C} \geq 2^{2(v-3)} + 3 \cdot 2^{v-7}(2^{v-3} - 1).$$

Proof. Since $n \equiv 4 \pmod{8}$, \mathbb{F}_{16} is a subfield of \mathbb{F}_{2^n} and we can choose W as the trace-zero plane in \mathbb{F}_{16} .⁶² Writing $\mathbb{F}_{16} = \mathbb{F}_2(\xi)$ with $\xi^4 + \xi + 1 = 0$ and $\omega = \xi^5$, we have $\mathbb{F}_4 = \mathbb{F}_2(\omega)$, $W = \{0, 1, \xi, \xi^2, \xi^4, \xi^5, \xi^8, \xi^{10}\} = \langle \xi, \omega \rangle$, and $\xi^2 + \xi = \omega = \omega(\omega^2 + \omega)$. This shows that W is of the type considered in Theorem 6.7 with $a = \omega$, $b = \xi$.⁶³ Further, from $W' = \{\delta(Z); Z \subset W\} = W$ and $\delta(W) = 1$ we find that the set of missing points of σ_W is $\{\delta(W)/\delta(Z)^3; Z \subset W\} = \{1, \xi^3, \xi^6, \xi^9, \xi^{12}\}$, the missing point of multiplicity 3 being 1.

⁶²The actual choice of W does not matter, since all planes in \mathbb{F}_{16} are rotated copies of each other (with factors $r \in \mathbb{F}_{16}^\times \subseteq \mathbb{F}_{2^n}^\times$) and hence in the same G -orbit. The subsequent proof, however, is only valid for the trace-zero plane, since it uses $W^2 = W$.

⁶³Strictly speaking, we should also check that $\text{Tr}(\omega) = \text{Tr}(\xi) = 0$ but this is trivial, since $\text{Tr}(x) = (n/4)\text{Tr}_{\mathbb{F}_{16}/\mathbb{F}_2}(x)$ for $x \in \mathbb{F}_{2^4} \subseteq \mathbb{F}_{2^n}$.

In what follows, since we have to deal with different traces simultaneously, we will adopt the simpler notation $\text{Tr}_{2^s}(x) = \text{Tr}_{\mathbb{F}_{2^n}/\mathbb{F}_{2^s}}(x) = x + x^{2^s} + x^{4^s} + \dots$ for $s \mid n$.

The collision space $(W^2)^\perp = W^\perp$ is easily seen to be $\{y \in \mathbb{F}_{2^n}; \text{Tr}_{16}(y) \in \mathbb{F}_2\}$ and intersects W in \mathbb{F}_2 .⁶⁴ This shows that 1 is the only missing point in $(W^2)^\perp$.

Now Theorem 6.6(iii) implies that \mathbf{C}_W has row sums 4 and 10 with corresponding frequencies $f_4 = 2^{n-4}$ and $f_{10} = 2^{n-4} - 1$. The $2^{n-4} \times (2^{n-3} - 1)$ submatrix $\mathbf{C}_W(I)$ formed by the rows of weight 4 has column sums ≤ 4 , since the supporting lines and planes in $\text{PG}(\mathbb{F}_{2^n}/W)$ (cf. Theorem 6.5(ii)) meet the affine subspace $\{T_i; i \in I\}$ in at most 2 points (resulting in a column sum $\leq 2 + 2 = 4$), respectively, at most 4 points (column sum $\leq 1 + 1 + 1 + 1 = 4$).⁶⁵ Hence the number of nonzero columns of $\mathbf{C}_W(I)$ must be at least 2^{n-4} , and we can take $m' = 2^{n-4}$ in Corollary 4 to conclude that

$$N_1 \geq 2^{n-4}(6 - 4) + 2^{n-4} = 3 \cdot 2^{n-4}.$$

This completes the proof. \square

Part (i) of our main theorem now follows from Theorem 8.1 and

$$3 \cdot 2^{v-7}(2^{v-3} - 1) > \frac{9}{8} \cdot \frac{(2^{v-4} - 1)(2^{v-3} - 1)}{3} = \frac{9}{8} \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2.$$

Remark 5. In the smallest case $v = 7$, in which \mathbb{F}_{2^n} coincides with the subfield \mathbb{F}_{16} , the maximum local net gain is equal to 3; cf. Example 5. Theorem 8.1 gives a lower bound for the maximum net gain at lengths $v = 7 + 8t$, $t = 1, 2, \dots$, which scales nicely with v and thus can be viewed as “anchored” at $v = 7$. Indeed, the proof of the theorem involves only computations in the subfield \mathbb{F}_{16} , no matter how large \mathbb{F}_{2^n} is. This point of view will become essential in the case $v \equiv 3 \pmod{8}$; see the next theorem. However, it should be noted that these observations only give lower bounds for the maximum net gain and that the actual maximum net gain can be substantially larger. For example, in the case $v = 15$ the maximum net gain is $924 > 3 \cdot 2^8 = 768$; cf. Table 1.

Theorem 8.2. *For packet lengths $v \equiv 3 \pmod{8}$, i.e., $n = v - 3 \equiv 0 \pmod{8}$, the maximum achievable local net gain N_1 in the general RRP satisfies $N_1 \geq 54 \cdot 2^{n-8} = 54 \cdot 2^{v-11}$, and hence the corresponding optimum subspace codes have size*

$$\#\mathcal{C} \geq 2^{2(v-3)} + 54 \cdot 2^{v-11}(2^{v-3} - 1).$$

Proof. Again taking W as the trace-zero plane in $\mathbb{F}_{16} \subset \mathbb{F}_{2^n}$, the proof remains the same as for Theorem 8.1 up to the point where the collision space is computed. The explicit formula for $(W^2)^\perp = W^\perp$ obtained earlier remains valid, but now the elements in \mathbb{F}_{16} have trace zero and hence are in $(W^2)^\perp$. In particular $(W^2)^\perp$ now contains all 5 missing points, and their geometric configuration must be taken into account. From $\xi^{12} = \xi^9 + \xi^6 + \xi^3 + 1$ it is clear that the 5 points form a projective basis of their hull $M = \mathbb{F}_{16}$ (i.e., are 5 points in general position). Giving the triple point homogeneous coordinates $(1 : 1 : 1 : 1)$, the corresponding linear $[7, 4]$ code

⁶⁴Here we use that $[\mathbb{F}_{2^n}/\mathbb{F}_{16}] = n/4$ is odd and hence $\text{Tr}(y) = \text{Tr}_{16}(y)$ for $y \in \mathbb{F}_{16} \subseteq \mathbb{F}_{2^n}$.

⁶⁵For the column of Type 4¹ the bound is trivial.

has generator matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

and weight distribution $A_0 = 1$, $A_2 = 6$, $A_4 = 5$, $A_6 = 4$. The corresponding row-sum spectrum is $m_{14} = 2^{n-7} - 1$, $m_{10} = 6 \cdot 2^{n-7}$, $m_6 = 5 \cdot 2^{n-7}$, $m_2 = 4 \cdot 2^{n-7}$. As before, let $\mathbf{C}_W(I)$ be the submatrix of \mathbf{C}_W formed by the rows with $r_i \leq 6$, i.e. $r_i = 2$ and $r_i = 6$. Our goal is to establish a lower bound m' on the number of nonzero columns of $\mathbf{C}_W(I)$, which is more difficult in this case.

First we note that the solids $\{T_i; i \in I\}$ are determined by $\text{Tr}(x) = 1$ (corresponding to the codewords with 1 or 3 nonzero entries among the first 4 coordinates) or $\text{Tr}(\xi^{3t}x) = 1$ for $1 \leq t \leq 4$ (corresponding to the codeword (1111000)). In $\mathbb{F}_{2^n}/M^\perp \cong \text{PG}(3, \mathbb{F}_2)$ these solids determine 9 points, the first 8 of which form an affine subspace (complement of the plane $\text{Tr}(x) = 0$).

Since $M = \mathbb{F}_{16}$, we have $(M^{1/2})^\perp = M^\perp = \mathbb{F}_{16}^\perp = \{x \in \mathbb{F}_{2^n}; \text{Tr}_{16}(x) = 0\}$ and can express the conditions in terms of $\text{Tr}_{16}(x)$. Using $\text{Tr}(x) = \text{Tr}_{\mathbb{F}_{16}/\mathbb{F}_2}(\text{Tr}_{16}(x))$, we find that the last point has equation $\text{Tr}_{16}(x) = 1$ and the 9 points are those with $\text{Tr}_{16}(x) \in (\mathbb{F}_{16} \setminus W) \cup \{1\}$.

Since $s_W(X) = X^8 + X^4 + X^2 + X$, $W' = W$, and $\delta(W) = 1$, we have from the proof of Theorem 6.5(i) that the entry of \mathbf{C}_W corresponding to $T = \langle W, x \rangle$ and $y \in W^\perp$ is the number of solutions of the equation

$$(19) \quad x^8 + x^4 + x^2 + x = y^2w^4 + yw \quad \text{in } W.$$

The above conditions on $\text{Tr}_{16}(x)$ translate into conditions on $\text{Tr}_{256}(x^8 + x^4 + x^2 + x)$; the first into $\text{Tr}_{16}(x^8 + x^4 + x^2 + x) = \text{Tr}(x) = 1$, which is equivalent to $\text{Tr}_{256}(x^8 + x^4 + x^2 + x) \in \{t \in \mathbb{F}_{256}; \text{Tr}_{\mathbb{F}_{256}/\mathbb{F}_{16}}(t) = 1\} = t_0 + \mathbb{F}_{16}$; and the second into $\text{Tr}_{256}(x^8 + x^4 + x^2 + x)^2 + \text{Tr}_{256}(x^8 + x^4 + x^2 + x) = \text{Tr}_{256}(x^{16} + x) = \text{Tr}_{16}(x) = 1$, i.e. $\text{Tr}_{256}(x^8 + x^4 + x^2 + x) \in \mathbb{F}_4 \setminus \mathbb{F}_2$.

On the other hand,

$$\text{Tr}_{256}(y^2w^4 + yw) = \text{Tr}_{256}(y)^2w^4 + \text{Tr}_{256}(y)w$$

depends only on $\text{Tr}_{256}(y)$ and hence is constant on cosets of $H_{256} = \{x \in \mathbb{F}_{2^n}; \text{Tr}_{256}(x) = 0\}$.

Putting the preceding observations together, we conclude that the total number of solutions of (19) with $T_i = \langle W, x \rangle$ varying over $i \in I$ is constant on cosets of H_{256} as well. This means that the frequencies in the column-sum spectrum of $\mathbf{C}_W(I)$ are obtained from those for $n = 8$ by scaling with 2^{n-8} . In particular, the number of nonzero columns of $\mathbf{C}_W(I)$ is $2^{n-8} \cdot t$, where t is the corresponding number for the case $n = 8$. For $n = 8$ we find by inspecting \mathbf{C}_W in Figure 1 that the 18×31 submatrix $\mathbf{C}_W(I)$ has 22 nonzero columns (16 columns of Type 1^4 and 6 columns of Type 1^2), resulting in $m' = 22$ and

$$N_1 \geq 8(6 - 2) + 22 = 54.^{66}$$

In the general case the bound then scales to $N_1 \geq 2^{n-8} \cdot 54$, as asserted. \square

⁶⁶In fact $N_1 = 54$, as shown in Section 7. With some effort the column types of $\mathbf{C}_W(I)$ can also be computed by hand. It turns out that the 16 columns corresponding to $y \in W^\perp$ with $\text{Tr}_{16}(y) = 1$ are those of type 1^4 , and the 6 columns corresponding to $y \in W \setminus \mathbb{F}_{16}$ (which have $\text{Tr}_{16}(y) = 0$) are those of type 1^2 .

$n \backslash \mu$	3	4	5	6	7
4	1	0	0	0	0
6	0	0	3	2	0
8	0	8	12	0	1
10	0	20	45	10	10
12	32	100	90	96	23
14	56	392	483	322	112
16	320	1360	2136	1376	269
18	1392	5388	8121	5546	1398
20	5616	21900	32550	21840	5475
22	22088	86240	131967	87362	21868

TABLE 2. Distribution of the number μ of missing points inside $(W_1^2)^\perp$

Again comparing the bound of Theorem 8.2 with the LMRD code bound, we obtain

$$54 \cdot 2^{v-11}(2^{v-3} - 1) > \frac{81}{64} \cdot \frac{(2^{v-4} - 1)(2^{v-3} - 1)}{3} = \frac{81}{64} \begin{bmatrix} v-3 \\ 2 \end{bmatrix}_2.$$

This proves Part (ii) of our main theorem.

The computational results presented in Table 1 show that the largest subspace codes obtained by solving the RRP exceed the LMRD code bound for all $v \in \{7, 8, \dots, 15\}$ except for $v = 8$ and $v = 10$. Although the margin is rather narrow for $v \in \{12, 14\}$, we make the following

Conjecture 1. *For any packet length $v \geq 7$, $v \notin \{8, 10\}$, the largest subspace codes that can be obtained by solving the RRP exceed the LMRD code bound and thus are better than the codes resulting from the echelon-Ferrers construction and its variants.*

By Theorems 8.1 and 8.2, Conjecture 1 is true for packet lengths $v \equiv 3 \pmod{4}$. For lengths $v \equiv 1 \pmod{4}$, which correspond to $n \equiv 2 \pmod{4}$, the following considerations provide strong evidence in support of Conjecture 1.

Inspecting the proof of Theorem 8.1, we see that the argument remains valid for $n \equiv 2 \pmod{4}$, provided there exists a plane W in $\text{PG}(\mathbb{F}_{2^n})$ which satisfies the conditions of Theorem 6.7 and has $\mu = 3$, i.e., there is a triple missing point inside $(W^2)^\perp$ and 4 missing points outside $(W^2)^\perp$. Using the explicit description of the missing points for the planes $W_1 = \langle 1, a, b \rangle$ considered in Theorems 6.6 and 6.7, it is easy to test the condition $P \in (W^2)^\perp$ and compute the number of planes W_1 with $\mu = 3$ for small values of n .⁶⁷

We have written a small SageMath worksheet for this job. The results are shown in Table 2. From the table we see that \mathbb{F}_{2^n} contains a plane W with a triple missing point and $\mu = 3$ for all $n \in \{12, 14, \dots, 22\}$. Moreover, the (shifted) frequency distribution of planes W_1 with μ missing points, normalized by the total number $(2^{n-2} - 1)/3$ of planes W_1 (cf. Theorem 6.7(ii)), seems to converge to the binomial distribution $(1/16, 4/16, 6/16, 4/16, 1/16)$.

⁶⁷Choosing $c = 1$ in Theorem 6.6(ii), the condition $P \in (W^2)^\perp$ reduces to $\text{Tr}(a_1/a_0^2) = 0$ or, somewhat easier to handle, $\text{Tr}((b + \omega a + x)^{-3}) = 0$; cf. Theorem 6.7(iii).

In particular, Conjecture 1 is also true for $n \in \{10, 14, 18, 22\}$, i.e., for packet lengths $v \in \{13, 17, 21, 25\}$.⁶⁸

It may be possible to prove Conjecture 1 for $v \equiv 1 \pmod{4}$ with the aid of character sums and the observations in Remark 4. The case of even v , however, seems to be much harder.

9. THE SIGNIFICANCE OF THE ASSOCIATED LINEAR CODE

From the discussion following Theorem 6.6 and the previous section we know already that the linear $[\mu, k]$ code $C = C_W$ associated to the multiset of missing points contained in the collision space $(W^2)^\perp$ plays an important role in computing the maximum net gain N_1 of the RRP relative to W . In this section we add further evidence to this by using C to express the bounds of Corollary 4 in terms of μ , k and showing that this refinement suffices to complete the solution of the RRP for $v \in \{14, 15\}$; this question was left open in Section 7.

The following lemma is implicit in the remarks following Theorem 6.6.

Lemma 9.1. *Given a plane W in $\text{PG}(\mathbb{F}_{2^v})$, let C be a binary linear $[\mu, k]$ code associated with the multiset \mathbf{m} of missing points of σ_W contained in $(W^2)^\perp$ (i.e., μ is the cardinality of \mathbf{m} and k the dimension of its hull M). Then the quantity $\sum_{r=0}^5 m_r(6-r)$ in Corollary 4 can be expressed in terms of the weight distribution A_0, \dots, A_μ of C as follows:*

$$\sum_{r=0}^5 m_r(6-r) = 2^{n-3-k} \times \sum_{i > (\mu+1)/2} (2i-1-\mu)A_i.^{69}$$

Proof. Just observe that r, i are related by $6-r = 6 - (\mu + 7 - 2i) = 2i - 1 - \mu$ and that the all-zero codeword of C , which corresponds to zero or more rows with sum $7 + \mu$, does not contribute to either side. \square

Planes W of the type considered in Theorem 6.7 are distinguished by the fact that the corresponding multiset \mathbf{m} is not a set; equivalently, the associated $[\mu, k]$ code C is not projective, or $d_{\text{Ham}}(C^\perp) = 2$.⁷⁰

Lemma 9.2. *For projective $[\mu, k]$ codes C , in the parameter range of interest to us, we have the following upper bounds on the “code sums” $\sum_{i > (\mu+1)/2} (2i-1-\mu)A_i$.*

$\mu \backslash k$	1	2	3	4	5	6	7
1	0						
2	0	1					
3	0	0	2				
4	0	0	3	7			
5	0	0	2	10	14		
6	0	0	3	9	20	38	
7	0	0	0	8	20	40	76

Moreover, these bounds are best possible.

⁶⁸The case $n = 10$, where no plane W_1 with $\mu = 3$ exists, is covered by Table 1.

⁶⁹In the case $\mu = 0$, where all rows of C_W have sum 7, the right-hand side should be interpreted as zero.

⁷⁰Note that $d_{\text{Ham}}(C^\perp) = 1$ is not possible, since by definition C has no all-zero coordinates.

$\gamma_{\mu,k}$	(μ, k)
0.625	(5, 4), (6, 5), (7, 5), (7, 6)
0.59375	(6, 6), (7, 7)
0.5625	(6, 4)
0.5	(7, 4)
0.4375	(4, 4), (5, 5)
0.375	(4, 3), (6, 3)
0.25	(2, 2), (3, 3), (5, 3)
0	otherwise

TABLE 3. Upper bounds on the normalized code sums of projective $[\mu, k]$ codes

For substituting the bounds into Lemma 9.1, it is convenient to normalize them by 2^{-k} . The resulting normalized upper bounds $\gamma_{\mu,k}$ are listed in the following table, in order of increasing strength.

Proof of Lemma 9.2. The entries in the diagonal of the table are the code sums obtained for the trivial $[\mu, \mu]$ codes. The zero entries are due to the fact that a projective $[\mu, k]$ code has $\mu \leq 2^k - 1$ and the simplex codes ($\mu = 2^k - 1$) have only codewords of weight 0 and $(\mu + 1)/2$.

The remaining cases are settled in an ad hoc fashion, using codes with a systematic generator matrix \mathbf{G} . The code sums yet relevant are

μ	code sum
4	$A_3 + 3A_4$
5	$2A_4 + 4A_5$
6	$A_4 + 3A_5 + 5A_6$
7	$2A_5 + 4A_6 + 6A_7$

$k = 3$. Viewing the columns of \mathbf{G} as points in the Fano plane $\text{PG}(2, \mathbb{F}_2)$, we have to consider 2 cases for $\mu = 4$ (a quadrangle and a line with one additional point, both of which have code sum 3) and one case for $\mu = 5, 6$ (having maximum weight 4 with $A_4 = 1$ and $A_4 = 3$ respectively).

$k = 4$. For $\mu = 5$ the even-weight subcode of \mathbb{F}_2^5 (with 5th column $(1111)^T$ in \mathbf{G}) is the unique code having code sum 10. For $\mu = 6$ there are 4 equivalence classes of codes with non-systematic parts $\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}^T$ and code sums 9, 9, 8, 8, respectively.⁷¹

For $\mu = 7$ there are 5 equivalence classes of codes with non-systematic parts $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$ and code sums 4, 8, 8, 6, 6, respectively.⁷²

$k = 5$. For $\mu = 6$ the even-weight subcode of \mathbb{F}_2^6 has code sum 20 is the only such code. For $\mu = 7$ there are 8 equivalence classes with non-systematic parts $\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}^T$, $\begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}^T$.

⁷¹The equivalence classes are best viewed as equivalence classes of the corresponding dual $[6, 2]$ codes. Note that C is projective iff C^\perp (which is not necessarily projective) has minimum weight ≥ 3 . In the case under consideration this restricts the column multiplicities of C^\perp to values ≤ 3 .

⁷²Now the column multiplicities of C^\perp are ≤ 2 .

$\begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}^\top$. The code sums are all ≤ 20 , with equality for the 6th and 8th equivalence class.

$k = 6$. Here we have only one case to consider, $\mu = 7$. The even-weight subcode of \mathbb{F}_2^7 has code sum 28 and is not “optimal” in this case. The codes with 7th column $(111110)^\top$, $(111100)^\top$ in their \mathbf{G} have code sum 40 and are the only such codes. \square

Theorem 9.3. *For $v \in \{14, 15\}$ the computed maximum local net gain in Table 1 (379 for $v = 14$, 924 for $v = 15$) represents the true maximum achievable net gain of the RRP.*

Proof. (i) $v = 14$ ($n = 11$). It suffices to show that any plane W in $\text{PG}(\mathbb{F}_{2048})$ that has at most 4 missing points in $(W^2)^\perp$ satisfies $N_1 < 379$. For $\mu \leq 4$ the maximum value of $\gamma_{\mu,k}$ in Table (3) is $\gamma_{4,4} = 0.4375$. From Corollary 4, Lemma 9.1 and the table, the maximum net gain relative to W satisfies

$$\begin{aligned} N_1 &\leq \sum_{r=0}^5 m_r(6-r) + 255 = 256\gamma_{\mu,k} + 255 \\ &< 1.4375 \times 256 = 368, \end{aligned}$$

as desired.

(ii) $v = 15$ ($n = 12$). Here we must show that any plane W in $\text{PG}(\mathbb{F}_{4096})$ that is not of the type considered in Theorem 6.7 has $N_1 < 924$. Since these planes are exactly those for which the associated $[\mu, k]$ code is projective, we can use the bound $\gamma_{\mu,k} \leq 0.625$ from Table (3) in Lemma 9.1. This gives

$$N_1 < 1.625 \times 512 = 832,$$

completing the proof of the theorem. \square

10. CONCLUSION

We conclude this paper with a list of open problems related to our work. Only the first problem has been discussed already (in Section 8).

Problem 1. Prove Conjecture 1, either partially for odd packet lengths $v \equiv 1 \pmod{4}$ or in full. *The case of odd v (even n) seems more accessible in view of the availability of planes with a triple missing point and the overwhelming evidence for the existence of such planes with $\mu = 3$, which would settle this part. The case of even v includes all cases where \mathbb{F}_{2^n} has prime degree over \mathbb{F}_2 and hence no nontrivial subfields. In this case an approach different from that in Theorems 8.1, 8.2 must be used, perhaps starting with an existence proof of planes W with a large code sum in their associated $[\mu, k]$ code (cf. Lemma 9.1) and using the lower bound in Corollary 4 with a suitable constant m' . Note that in terms of the size of \mathbf{C}_W , the threshold for the local net gain set by the LMRD code bound is $\frac{2^{n-1}-1}{3} \approx \frac{4}{3} \times (2^{n-3} - 1)$.*

Problem 2. Improve the expurgation-augmentation method for small packet lengths v . *Although our method represents an asymptotic improvement of the known constructions of $(v, M, 4; 3)_2$ codes for $v \mapsto \infty$, it is much inferior to the group-invariant computational constructions in [10] for lengths $v \in \{8, 9, 10, 11\}$. To some extent this can be remedied through adding a further computational extension step by planes meeting S in a line (cf. the remarks in Section 7), but the results remain inferior to [10], and with increasing length the method soon becomes infeasible.*

To overcome this problem, an algebraic description of the free planes relative to an optimal solution of the RRP (or a suitable subcode thereof, which avoids “colliding planes”) would be desirable. Another approach, which for $v = 8$ at least yields some improvement,⁷³ is to compute, relative to the expurgated lifted Gabidulin code, the set of all free planes meeting S in a point and use a suitable maximum-clique algorithm to find the absolutely largest extension of the expurgated code by such planes. We have determined experimentally that usually there are indeed additional free planes (corresponding to non-standard rearrangements of the free lines into new planes). However, including those planes in the optimization problems destroys its rotation-invariance and makes it much more computationally expensive. Again an algebraic description of the set of all free planes may help to overcome this problem.

Problem 3. Use the expurgation-augmentation method with other LMRD codes or subsets thereof. The Gabidulin codes \mathcal{G}_W considered in this paper are not the only MRD codes with these parameters, provided that $v \geq 7$. Therefore the question arises whether one can adapt the expurgation-augmentation method for use with other LMRD codes and, if so, what the maximum sizes of the corresponding modified subspace codes will be. Although it is not directly related to this question, the following observation made during the preparation of [31] may be of interest in this regard: One of the five isomorphism types of optimal $(6, 77, 4; 3)_2$ codes, named Type B in [31], contains a set of 16 planes disjoint from $S = \{\mathbf{0}\} \times \mathbb{F}_2^3$ at mutual subspace distance 4. This set corresponds to a 4-dimensional constant-rank-two subspace of $\mathbb{F}_2^{3 \times 3}$ of the type discovered by Beasley [1]. Since Gabidulin codes in $\mathbb{F}_2^{3 \times 3}$ do not contain such a “Beasley code”, we have that $(6, 77, 4; 3)_2$ codes of Type B cannot be obtained by ordinary expurgation-augmentation as considered in this paper.

Problem 4. Generalize the expurgation-augmentation method to subspace codes of constant dimension $k > 3$. As an example we consider the smallest length $v = 8$, for which this problem is meaningful. For $v = 8$ there are two cases with $k > 3$, where $A_2(v, d; k)$ is unknown, viz. $(v, d; k) = (8, 4; 4)$ and $(8, 6; 4)$. In the first case the corresponding Gabidulin code \mathcal{G} provides a set of 2^{12} solids, which are disjoint from $S = \{\mathbf{0}\} \times \mathbb{F}_2^4$ and cover each plane in $\text{PG}(\mathbb{F}_2^8)$ disjoint from S exactly once. Since solids disjoint from S contain 15 such planes, while solids meeting S in a point contain only 8 such planes, it should be possible—at least in principle—to rearrange the planes in some subset of \mathcal{G} into new solids meeting S in a point and thereby increase the code size significantly. However, the details seem a lot more involved than in the case $k = 3$, and suitable subsets of \mathcal{G} have yet to be found. The same remark applies to the other parameter triple $(8, 6; 4)$, in which the corresponding Gabidulin code \mathcal{G} consists only of 2^8 solids disjoint from S and covers each line disjoint from S exactly once.⁷⁴

REFERENCES

- [1] L. B. Beasley, Spaces of rank-2 matrices over $\text{GF}(2)$, *Electronic Journal of Linear Algebra*, **5** (1999), 11–18.

⁷³The largest $(8, M, 4; 3)_2$ code obtained has size $M = 1286$, compared with 1259 in Section 7 and 1326 in [10].

⁷⁴The currently best lower bound in this case is still the rather trivial $A_2(8, 6; 4) \geq 257$, coming from $\mathcal{G} \cup \{S\}$, and it may well give the true result.

- [2] E. Ben-Sasson, T. Etzion, A. Gabizon and N. Raviv, Subspace polynomials and cyclic subspace codes, 2014, Preprint arXiv:1404.7739 [cs.IT].
- [3] E. Ben-Sasson and S. Kopparty, Affine dispersers from subspace polynomials, *SIAM Journal on Computing*, **41** (2012), 880–914.
- [4] E. Ben-Sasson, S. Kopparty and J. Radhakrishnan, Subspace polynomials and limits to list decoding of ReedSolomon codes, *IEEE Transactions on Information Theory*, **56** (2010), 113–120.
- [5] E. R. Berlekamp, *Algebraic coding theory*, McGraw-Hill, 1968.
- [6] Jan De Beule and Leo Storme (eds.), *Current Research Topics in Galois Geometry*, Nova Science Publishers, 2011.
- [7] A. Beutelspacher, Partial spreads in finite projective spaces and partial designs, *Mathematische Zeitschrift*, **145** (1975), 211–230, Corrigendum, *ibid.* 147:303, 1976.
- [8] S. R. Blackburn and T. Etzion, The asymptotic behavior of Grassmannian codes, *IEEE Transactions on Information Theory*, **58** (2012), 6605–6609.
- [9] M. Braun, T. Etzion, P. R. J. Östergård, A. Vardy and A. Wassermann, Existence of q -analogs of Steiner systems, 2013, Preprint arXiv:1304.1462 [math.CO].
- [10] M. Braun, P. Östergård and A. Wassermann, New lower bounds for binary constant dimension subspace codes, 2015, Preprint.
- [11] M. Braun and J. Reichelt, q -analogs of packing designs, *Journal of Combinatorial Designs*, **22** (2014), 306–321, Preprint arXiv:1212.4614 [math.CO].
- [12] H. E. Campbell and D. L. Wehlau, *Modular Invariant Theory*, Springer-Verlag, 2011.
- [13] Q. Cheng, S. Gao and D. Wan, Constructing high order elements through subspace solynomials, in *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms (SODA'12)*, Society for Industrial and Applied Mathematics, 2012, 1457–1463.
- [14] A. Cossidente and F. Pavese, On subspace codes, *Designs, Codes and Cryptography*, Electronically published on Oct 30, 2014.
- [15] P. Dembowski, *Finite Geometries*, Springer-Verlag, 1968, Classics in Mathematics Series, 1997.
- [16] H. Derksen and G. Kemper, *Computational Invariant Theory*, Springer-Verlag, 2002.
- [17] L. E. Dickson, A fundamental system of invariants of the general modular linear group with a solution of the form problem, *Transactions of the American Mathematical Society*, **12** (1911), 75–98.
- [18] J. Eisfeld and L. Storme, (partial) t -spreads and minimal t -covers in finite projective spaces, 2000, Lecture notes, Ghent University.
- [19] S. El-Zanati, H. Jordon, G. Seelinger, P. Sissokho and L. Spence, The maximum size of a partial 3-spread in a finite vector space over $\text{GF}(2)$, *Designs, Codes and Cryptography*, **54** (2010), 101–107.
- [20] T. Etzion, Problems on q -analogs in coding theory, 2013, Preprint arXiv:1305.6126 [cs.IT].
- [21] T. Etzion and N. Silberstein, Error-correcting codes in projective spaces via rank-metric codes and Ferrers diagrams, *IEEE Transactions on Information Theory*, **55** (2009), 2909–2919.
- [22] T. Etzion and N. Silberstein, Codes and designs related to lifted MRD codes, *IEEE Transactions on Information Theory*, **59** (2013), 1004–1017, Erratum *ibid.* 59(7):4730, 2013.
- [23] T. Etzion and L. Storme, Galois geometries and coding theory, *Designs, Codes and Cryptography*, **78** (2016), 311–350.
- [24] T. Etzion and A. Vardy, Error-correcting codes in projective space, *IEEE Transactions on Information Theory*, **57** (2011), 1165–1173.
- [25] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, 1979.
- [26] X. Guang and Z. Zhang, *Linear Network Error Correction Coding*, SpringerBriefs in Computer Science, Springer-Verlag, 2014.
- [27] D. Heinlein, M. Kiermaier, S. Kurz and A. Wassermann, Tables of subspace codes, 2016, Preprint arXiv:1601.02864 [math.CO].
- [28] J. W. P. Hirschfeld, *Projective Geometries over Finite Fields*, 2nd edition, Oxford University Press, 1998.
- [29] T. Honold and M. Kiermaier, On putative q -analogues of the Fano plane and related combinatorial structures, 2015, Preprint arXiv:1504.06688 [math.CO].
- [30] T. Honold, M. Kiermaier and S. Kurz, Constructions and bounds for mixed-dimension subspace codes, 2015, Preprint arXiv:1512.06660 [math.CO].

- [31] T. Honold, M. Kiermaier and S. Kurz, Optimal binary subspace codes of length 6, constant dimension 3 and minimum subspace distance 4, in *Topics in Finite Fields. 11th International Conference on Finite Fields and their Applications, July 22–26, 2013, Magdeburg, Germany* (eds. G. Kyureghyan, G. L. Mullen and A. Pott), vol. 632 of Contemporary Mathematics, American Mathematical Society, 2015, 157–176, Preprint arXiv:1311.0464 [math.CO].
- [32] R. Koetter and F. Kschischang, Coding for errors and erasures in random network coding, *IEEE Transactions on Information Theory*, **54** (2008), 3579–3591.
- [33] A. Kohnert and S. Kurz, Construction of large constant dimension codes with a prescribed minimum distance, in *Mathematical Methods in Computer Science. Essays in Memory of Thomas Beth* (eds. J. Calmet, W. Geiselmann and J. Müller-Quade), no. 5393 in Lecture Notes in Computer Science, Springer-Verlag, 2008, 31–42.
- [34] F. R. Kschischang, An introduction to network coding, in *Network Coding: Fundamentals and Applications* (eds. M. Médard and A. Sprintson), Elsevier Science Publishers, 2012, chapter 1, 1–37.
- [35] S. Kurz, Improved upper bounds for partial spreads, 2015, Preprint arXiv:1512.04297 [math.CO].
- [36] H. Liu and T. Honold, Poster: A new approach to the main problem of subspace coding, in *9th International Conference on Communications and Networking in China (ChinaCom 2014, Maoming, China, Aug. 14–16)*, 2014, 676–677, Full paper available as arXiv:1408.1181 [math.CO].
- [37] K. Metsch, Bose-Burton type theorems for finite projective, affine and polar spaces, in *Surveys in Combinatorics, 1999* (eds. J. D. Lamb and D. A. Preece), no. 267 in London Mathematical Society Lecture Note Series, Cambridge University Press, 1999, 137–166.
- [38] E. H. Moore, A two-fold generalization of Fermat’s theorem, *Bulletin of the American Mathematical Society*, **2** (1896), 189–199.
- [39] O. Ore, On a special class of polynomials, *Transactions of the American Mathematical Society*, **35** (1933), 559–584, Corrigendum *ibid.* 36(2):275, 1934.
- [40] N. Silberstein and A.-L. Trautmann, Subspace codes based on graph matchings, Ferrers diagrams, and pending blocks, *IEEE Transactions on Information Theory*, **61** (2015), 3937–3953.
- [41] D. Silva, F. Kschischang and R. Koetter, A rank-metric approach to error control in random network coding, *IEEE Transactions on Information Theory*, **54** (2008), 3951–3967.
- [42] D. Silva, F. Kschischang and R. Koetter, Communication over finite-field matrix channels, *IEEE Transactions on Information Theory*, **56** (2010), 1296–1306.
- [43] L. Smith, Polynomial invariants of finite groups. a survey of recent developments, *Bulletin of the American Mathematical Society*, **34** (1997), 211–250.
- [44] L. Storme and A. Nakić, On the extendability of particular classes of constant dimension codes, *Designs, Codes and Cryptography*, Electronically published on Aug 2, 2015.
- [45] A.-L. Trautmann and J. Rosenthal, New improvements on the Echelon-Ferrers construction, in *Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS 2010)* (ed. A. Edelmayer), Budapest, Hungary, 2010, 405–408, Reprint arXiv:1110.2417 [cs.IT].
- [46] H. Wang, C. Xing and R. Safavi-Naini, Linear authentication codes: Bounds and constructions, *IEEE Transactions on Information Theory*, **49** (2003), 866–872.
- [47] C. Wilkerson, A primer on the Dickson invariants, in *Proceedings of the Northwestern Homotopy Theory Conference (Evanston, Ill., 1982)*, vol. 19 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1983, 421–434.
- [48] S.-T. Xia and F.-W. Fu, Johnson type bounds on constant dimension codes, *Designs, Codes and Cryptography*, **50** (2009), 163–172.

E-mail address: somnus@zju.edu.cn

E-mail address: honold@zju.edu.cn

E-mail address: liuhaiteng@zju.edu.cn